SOME CONTEXTUAL EFFECTS ON THE PERCEPTION OF SYNTHETIC VOWELS

By

CARL LOUIS THOMPSON

A DISSERTATION PRESENTED TO THE GRADUATE COUNCIL OF
THE UNIVERSITY OF FLORIDA
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

December, 1965

ACKNOWLEDGMENTS

Sincere appreciation is expressed to Dr. Harry Hollien for his continued stimulation throughout an extensive and rewarding academic association, and particularly for his able supervision of the author's graduate program and research endeavors.

The author also wishes to acknowledge the teaching and counseling, and particularly criticisms, suggestions, and encouragement during the preparation of the manuscript, of Drs. Donald Dew, Richard Anderson, McKenzie Buck, and George Singleton.

The sincere thanks of the author are also extended to his wife, Shirley Kemper Thompson.

TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS	ii
LIST OF TABLES	iv
LIST OF FIGURES	v
REVIEW OF THE LITERATURE AND PURPOSE	1
Introduction	1
Purpose	5
PROCEDURE	7
Overview	7
Preparation of vowel samples	8
Human vowels	10
Synthetic vowels	16
Inter-vowel periods	19
Preparation of experimental tapes	19
Listener selection	20
Experimental procedure	20
Data reduction and analysis	21
RESULTS	23
DISCUSSION	49
Direction	50
Inter-vowel interval	50
Vowel ambiguity	51
SUMMARY AND CONCLUSIONS	53
BIBLIOGRAPHY	55
APPENDIX A	57
APPENDIX B	59

LIST OF TABLES

Table		Page
1.	Fundamental frequencies and mean ratings for the five vowel samples used as human productions.	15
2.	Formant frequencies, bandwidths, and F_1/F_2 ratios for nine synthetic stimuli used as initial and final vowels.	17
3.	Numbers of responses in each of five categories for each initial vowel, following vowel combination .	24
4.	Proportions of greater responses for each initial vowel-following vowel combination	27
5.	Vowel pair difference categories for the five initial vowels with each following vowel.	37
6.	Summary of an analysis of variance for the main effects, and interactions of initial vowel source, initial vowel, inter-vowel period, and following vowel	40
7.	Chi square values for Lv>Fv vs. Lv<Fv for each Fv.	47

LIST OF FIGURES

Figure		Page
1.	Fairbanks and Grubb (1961) frequency areas of formants one and two for preferred vowel samples.	9
2.	Frequencies of formants one and two for nine vowels used in the present study superimposed upon the Fairbanks and Grubb data for samples of five vowels. .	11
3a.	Percentage of greater responses as a function of the initial vowel for the following vowel / ɪ /	28
3b.	Percentage of greater responses as a function of the initial vowel for the following vowel X_1	29
3c.	Percentage of greater responses as a function of the initial vowel for the following vowel X_2	30
3d.	Percentage of greater responses as a function of the initial vowel for the following vowel / ɛ /	31
3e.	Percentage of greater responses as a function of the initial vowel for the following vowel X_3	32
3f.	Percentage of greater responses as a function of the initial vowel for the following vowel X_4	33
3g.	Percentage of greater responses as a function of the initial vowel for the following vowel /ʌ /.	34
4.	Summary of data for all vowel pairs showing the mean change in percentage of greater responses for seven categories of initial vowel-following vowel difference	36
5.	Observed and expected percentages of greater responses for each initial vowel at two inter-vowel intervals. .	42
6.	Summary of the data for all vowel pairs showing the mean differences in the percentage of greater responses between two inter-vowel intervals, .1 and .5 seconds, for seven categories of initial vowel-following vowel difference	43

7. Percentages of greater responses for each following vowel at each of two inter-vowel intervals. 46

REVIEW OF THE LITERATURE AND PURPOSE

Introduction

Acoustic energy such as that comprising the human vowel is usually specified in terms of fundamental frequency, intensity, and spectral composition (Black, 1939; Potter, Kopp, and Green, 1947; Stevens and Davis, 1938). Moreover, certain aspects of the spectral composition, i.e., formant frequencies, are cited (DeLattre, *et al.*, 1952; Dunn, 1950; Fairbanks and Grubb, 1961; Peterson, 1952; Potter and Steinberg, 1950) as the primary determinants of vowel quality. Specifically, vowel formants are frequency regions of energy concentration which are generally attributed to nonlinearities in the transfer function of the supralaryngeal cavities (Fant, 1952; Lewis, 1936; Lewis and Tuthill, 1940; Peterson and Barney, 1952; Potter and Steinberg, 1950; van den Berg, 1955; Stevens and House, 1961).

The studies cited above have shown that there is a fundamental relationship between the vowels produced and the center frequencies of the lower two formants (designated as F_1 and F_2). When the measured F_1 versus F_2 values for spoken vowels are displayed in a scattergram, it is generally observed that large proportions of the samples of each vowel fall in relatively small areas. However, when the data points for each vowel are enclosed by smooth curves, there are usually small areas of overlap which contain points for two or more vowels.

Variability in formant frequency patterns has been attributed to two types of sources. First there are differences which can be observed among a given speaker's replications of a vowel. Black's (1939) results indicated that one main cause of this type of variability is the consonantal context in which the vowel is produced. He utilized Fourier analysis to measure the spectral composition of vowel samples spoken by a single subject who produced similar vowels, both in identical and varying consonantal contexts. Considerably more variability was found for the productions in which context was systematically changed. The data of House, Stevens, and Fujisaki (1960) are in general agreement with those of Black. Using an analysis by synthesis technique, they measured formant frequencies of three adult males producing several vowels in a variety of consonantal contexts. Each vowel was pronounced as the stressed syllable in a bisyllabic nonsense "word." They found that their vowel formant frequencies differed systematically from published data derived from more restricted consonantal environments. While they do not report the analysis of such restricted environments, it seems probable that vowels in a bisyllabic context would add substantially to the overall variability of a study which included both types of syllables.

A second source of variability occurs in different speakers' productions of a vowel in a single consonantal context. Peterson and Barney (1952) have reported F_1/F_2 data obtained on seventy-six speakers producing two samples of each of ten vowels in an "h_d" context. Of these speakers, thirty-three were men, eighteen were women, and fifteen were children. Inter-subject differences were

highly significant, indicating that for at least some vowels, there are nonrandom variations in the formant frequencies used by the different individuals. Despite this variability, a graphic presentation of their F_1/F_2 frequencies shows relatively compact clusters for productions of each vowel.

The Peterson-Barney recordings were also presented to a group of seventy listeners for identification. When only those samples which met a criterion of correct identification by 100 per cent of the listening group were included, both the variability and vowel overlap were reduced, but were still "greater...than might be expected." They also state that a portion of the variability is due to differences in the three groups of speakers--men, women, and children. Although the mean formant frequencies are listed for each group; data regarding the portion of the overall variability due to the mixing of these three subgroups is not specified.

In the studies discussed above, there is evidence that vowel samples, which were identified as the same phoneme, may vary considerably in their spectral characteristics. In one study, that of Peterson and Barney (1952), it has been shown specifically that correct (i.e., 100 per cent) identification by a relatively large group of listeners, is possible despite overlaping formant frequencies. It seems evident, therefore, that some factor or factors--in addition to the F_1/F_2 characteristics of the vowel productions--must be responsible for the correct identification of items in these overlaping areas. One parameter which could affect a listener's identification of a vowel sample is vowel context, specifically the relationships of the F_1/F_2 frequencies of the vowel being identified

to those of other vowels which have immediately preceded that sample.

Ladefoged and Broadbent (1957) have demonstrated that such a contextual effect can occur. They produced synthetically six examples of the sentence, "Please say what this word is," with different sets of formant frequencies for the vowels in each of the six words. The carrier phrase was followed by one of the five vowels in a "p_t" consonantal context. They report that the identifications of these vowels changed as a function of the formant frequencies used in the six words in the carrier sentence.

Data on a seemingly different contextual effect are reported by Fry (1964) in his description of a study actually designed to investigate other relationships. In this research, judges were asked to assign two formant productions to one of the three categories. An order effect was found which caused certain of the vowels to be classified one way when preceded by a specific vowel and as a different phoneme when preceded by another. The effect was one of contrast, that is, a production tended to be identified as being a vowel which was less like the preceding vowel, than it would have been were both presented separately. In his discussion of this effect Fry cites the previously discussed work of Ladefoged and Broadbent (1957) as well as these data, as supporting Joos' (1948) statement, "on first meeting a person, the listener hears a few vowel phones, and on the basis of this small but apparently sufficient evidence he swiftly constructs a fairly complete vowel pattern to serve as a background (coordinate system) upon which he correctly locates new phones as fast as he hears them."

Fry, et al. (1962) considered the contrast effect in greater detail. Specifically, in regard to the contextual relationship of one vowel to another, they state:

> these results...also support the view...that in dealing with vowels uttered by a particular speaker, listeners rapidly form an appropriate reference frame against which they judge the quality of, and identify the sounds which occur. The reference frame is readily changed when utterances from another speaker are received and it is clearly dependent on judgements of the relations between vowel qualities.

In summary, the above discussion indicates that a significant contextual effect of one vowel upon another may exist; however, little if any information is available regarding the magnitude of the effect or with respect to what factors may influence it. As an example, Fry has provided some data relevant to the direction of the shift, but not concerning its relative magnitude. Moreover, it has not been indicated whether the effect varies as a function of the degree of the physical difference between the affecting and effected vowels.

Purpose

The purpose of this investigation is to study changes in the recognition of vowels that result from the affect of immediately preceding vowels. The three specific sub-questions are as follows:

1. Does such an effect exist; is it consistent?
2. If so, does its direction vary as a function of the relationship of the F_1/F_2 ratios of the affecting and effected vowels?
3. What is the relative magnitude of the effect?

Two additional related questions are asked:

1. Do listeners respond differently if the first vowel of a pair is a human production rather than one that is synthetic?
2. Does the strength of the effect vary as a function of the duration of the interval between affecting and effected vowels?

PROCEDURE

Overview

The purpose of this investigation was to determine if the identification of vowel quality in selected vowel-like sounds is affected by the characteristics of an immediately preceding vowel. In order to accomplish this, recordings of pairs of vowels were presented to twenty-five listeners who were instructed to identify the second of each pair of stimuli.

The seven judged stimuli were generated synthetically. Three were intended to be good examples of the vowels / ɪ /, / ɛ /, and / ʌ /. The formant frequencies of the other four stimuli were intended to create vowels which would be ambiguous. Two had formant frequencies between those of the synthetic vowels / ɪ / and / ɛ /; the others between / ɛ / and / ʌ /.

The five initial, or effecting vowels, were /i /, / ɪ /, / ɛ /, / ʌ /, and / ɑ /; both human and synthetic samples of each were used. This was done so that the applicability of previous and future research with synthetic vowels could be evaluated. It was felt that if similar effects are noted with both human and synthetic productions, the generality of other research using only synthetic samples would be confirmed.

In order to gain information as to whether the effect varies temporally, two inter-vowel periods, .1 and .5 seconds, were used. Each possible combination of the two sets (human and synthetic) of

five initial vowels, the two inter-vowel periods, and the seven following vowels, was presented twice for a total of 280 test items.

Data were analyzed for the affects of the initial vowel of each pair upon the identification of the second member of that pair. Further analysis was carried out in an attempt to discover any differences due to the two inter-vowel periods, or between the effects of synthetic and human vowels.

Preparation of vowel samples

The vowels /i /, /ɪ /, /ɛ /, /ʌ /, and /ɑ / were selected because they form a reasonably large set of phonemes in which there is a consistent decrease in the first formant frequency associated with an increase in second formant frequency. The specific formants used to produce the synthetic vowels were taken from data reported by Fairbanks and Grubb (1961). These authors list mean frequencies for the first three formants of preferred samples of nine American English vowels. Their samples, taken from steady state productions, were highly selected for "representativeness and identifiability." Figure 1 shows the frequency areas which they report for the first and second formants of these productions. The pattern of formant differences in the five vowels selected for the present study is readily seen in this presentation. For these five vowels, their formant data show an increase in formant one and a decrease in formant two which is almost linear when plotted logarithmically. The formant frequencies of the four intermediate vowels were selected with respect to the F_1/F_2 frequencies of the vowels / ɪ /, / ɛ /, and /ʌ / described above; those of X_1 and X_2 to create ambiguous vowels which fall between / ɪ / and / ɛ /; those of X_3 and

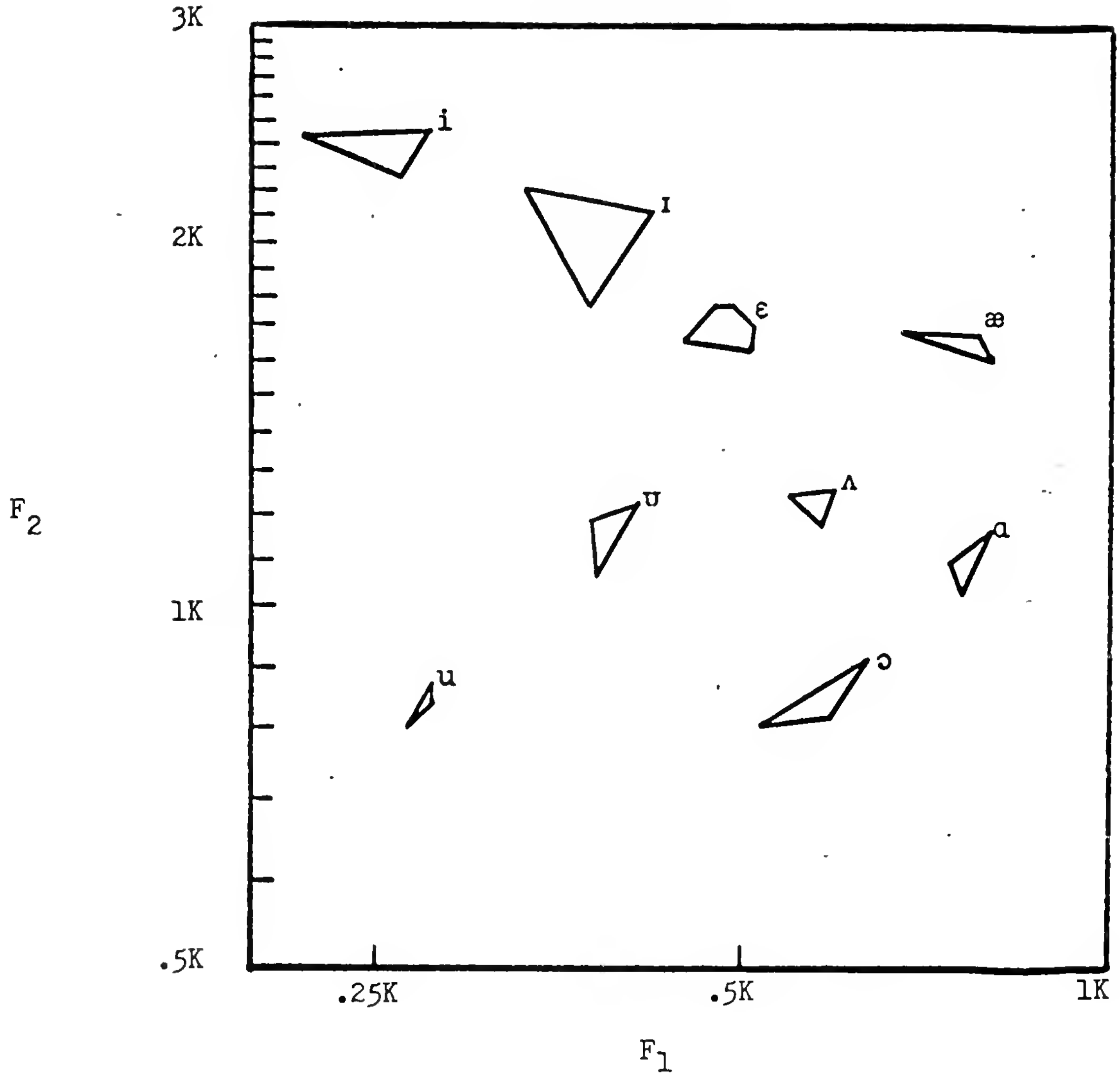

Figure 1. Fairbanks and Grubb (1961) frequency areas of formants one and two for preferred vowel samples. Values are in cps.

X_4 to create vowels which fall between /ɛ / and /ʌ /. From Figure 2, it can be seen that the formant intersects for X_1 and X_2 trisect a line drawn between the intersects of /ɪ / and /ɛ / while those of X_3 and X_4 trisect the line between /ɛ / and /ʌ /. Actually, however, frequencies of these formants were derived mathematically from the values found in Table 2 for the formants of the vowels /ɪ /, /ɛ /, and /ʌ /. For example, the logarithms of the first formants of /ɪ / and /ɛ / are 2.5798 and 2.6902, respectively. The 410 cycles/per/ second value used for the first formant of X_1 was obtained by adding one-third of the difference between /ɪ / and /ɛ / (.0368) to the value for /ɪ / and converting this value to frequency. Similarly the frequency for formant two of X_1 was lower than that of second formant of the vowel /ɪ / by an amount equal to one-third the logarithmic difference between the second formants of /ɪ / and /ɛ /. An identical process was used to obtain the formant one and formant two frequencies for the other three intermediate vowels.

Human vowels.--In order to obtain the human vowel productions, the procedure by Fairbanks and Grubb (1961) used to obtain "preferred vowel samples," was replicated as closely as possible. Since the present study is not concerned with inter-speaker differences in the effect being studied, only one speaker was used. He is a member of the faculty of the Communication Sciences Laboratory at the University of Florida, a habitual user of the General American dialect, and an experienced phonetician familiar with the Fairbanks and Grubb (1961) procedure. He was asked to produce clearly identifiable samples of the five vowels /i /, /ɪ /, /ɛ /, /ʌ /, and /ɑ /.

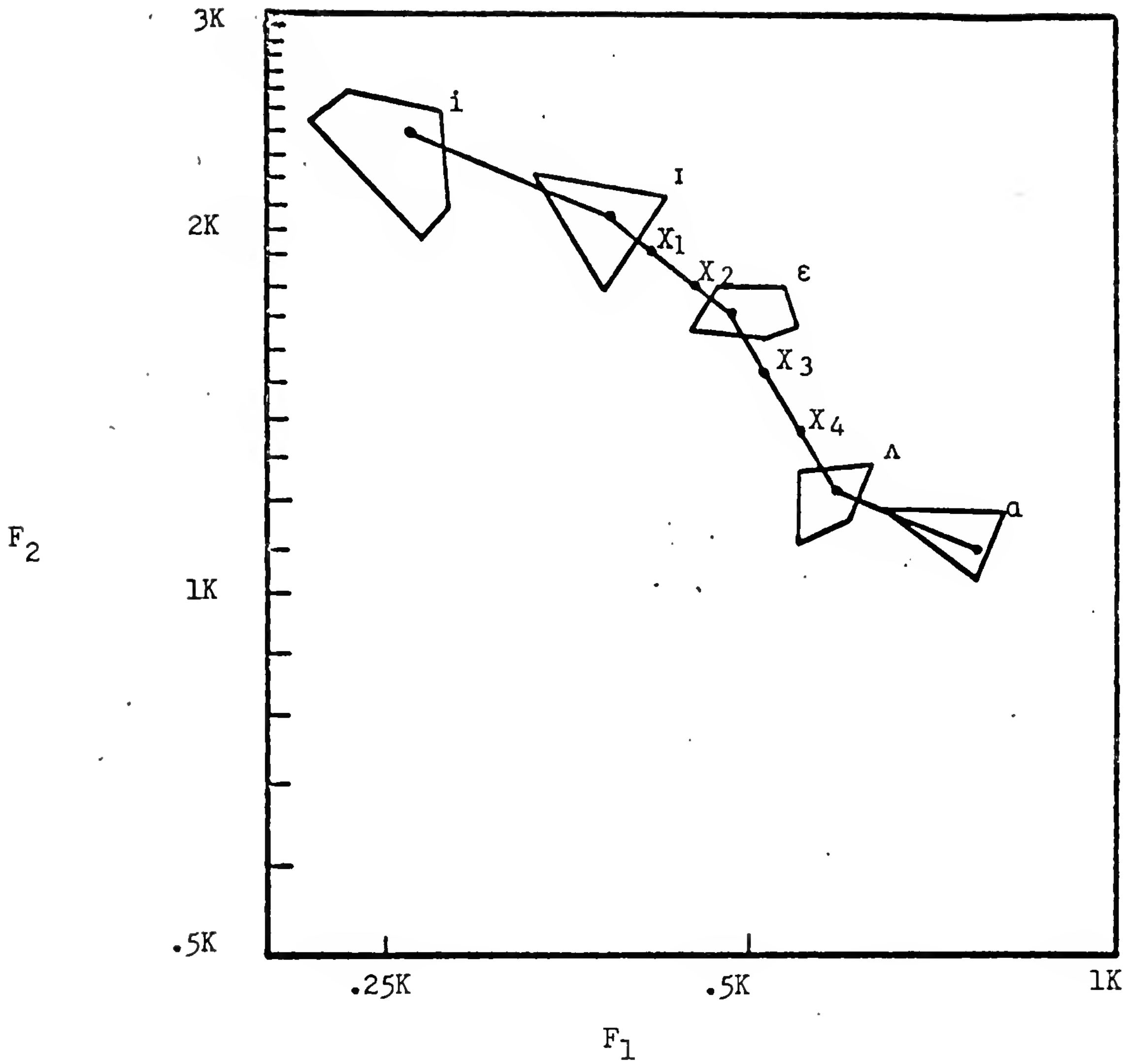

Figure 2. Frequencies of formants one and two for nine vowels used in the present study superimposed upon the Fairbanks and Grubb data for samples of five vowels. Values are in cps.

Recordings were made in an IAC-403A sound-treated room situated within a custom-built acoustically-treated area. The equipment included an Altec M-20 microphone and matching power supply coupled to an Ampex 350 full-track tape recorder. All recordings were made at 15 inches/per/second on 30 inch tape loops; sample duration and onset-offset time were controlled by a Grason-Stadler 829D electronic switch and 471-1 interval timer. The recorded samples had the same duration as those used by Fairbanks and Grubb (1961), i.e., 318 milliseconds ± 1 millisecond; the rise and decay times were 25 milliseconds. Black (1939) and Tiffany (1953) have shown that there are differences in the average durations of these vowels in speech. However, it was felt that any naturalness which would be gained by varying vowel duration appropriately would be more than offset by the difficulties in determining what portion of the experimental effect might be due to such a procedural variation. The recordings were replayed for the speaker's evaluation through a Marantz Model 7 preamplifier and 8B amplifier coupled to an Acoustic Research AR-3 speaker system.

In order to assist the speaker in producing the vowels at the desired fundamental frequency, a reference tone was provided. This was produced by a Hewlett Packard 202-CR oscillator set at 130 cps, and checked periodically with a Hewlett Packard 552-B electronic counter. The oscillator drove two Telephonic TDH-39 earphones, one for the speaker and the second for the experimenter.

Ten productions of each of the vowels were obtained in the following manner. The speaker listened to the reference tone and when ready, produced a sustained sample of one of the vowels. When

the experimenter felt that the speaker was producing the desired vowel, and that his frequency was matched to the reference tone, he triggered the timer which allowed a segment of that vowel to be recorded. The speaker and experimenter then evaluated the recording and, if they found it acceptable, it was retained. Two acceptable productions of each vowel were recorded in this manner. This sequence was repeated five times to obtain a total of fifty samples, ten each of the five vowels. Each vowel was produced at a "comfortable" loudness level.

In order to select the human vowel samples for the experimental procedure, a vowel selection tape consisting of 150 test items and twenty practice items was constructed. Each of the fifty items obtained was presented three times in random order. The intensity of the items was equalized to within ± 1 dB by adjusting the recorder gain control before each item was dubbed. The judging interval between vowels was a relatively long period (eight seconds) which was specified in an attempt to avoid any affect the preceding item might have on the judges' responses. An additional four-second pause, to provide listener orientation to the task, separated each group of five items. This corresponded to the spacing of the response form and was intended to compensate for the lack of item identification on the tape. Such identification was not used in order to avoid the development of a contextual framework by the judges. Each tape loop was dubbed from an Ampex 350 tape deck onto a Magnecord M-90 tape recorder, which was also used to replay the vowel selection tape for the judges. The listening environment, amplifier, and speaker system have been described previously.

Five judges who have had considerable experience in evaluating steady state vowel quality were selected from among the faculty and graduate students at the Communication Sciences Laboratory and the Speech Department at the University of Florida.

By means of a forced-choice technique, the judges identified each stimulus as one of a closed set of five vowels and then evaluated it on a nine point "quality" scale. They were instructed to base their evaluations on an estimate of that particular sample's representativeness as an example of that vowel as it most frequently occurs in the General American Dialect. Scores of one through three were used to indicate varying degrees of certainty or uncertainty that the vowel heard was not one of the specified five vowels, but would be better described as some other vowel. Ratings of four through nine were used to indicate the degree of success with which the sample represented the vowel as identified. The sample of each vowel that received 100 per cent correct identification and the highest average judged score was selected for fundamental frequency measurement. The average scores for the five vowels selected may be seen in Table 1.

As stated, the tape loops of the sample of each vowel having the highest score were evaluated for fundamental frequency. Each tape loop was reproduced on an Ampex 354 tape recorder coupled to an Allison 420 band-pass filter set to pass a one-third octave band of frequencies centered at 130 cps. The filter output was fed to a Marantz Model 7 preamplifier and 8B amplifier driving a speaker of appropriate impedance. A second input to the preamplifier consisted of a Hewlett Packard 202-CR oscillator. The tape loop was played repeatedly and the frequency of the oscillator varied about 130 cps

Table 1. Fundamental frequencies and mean ratings for the five vowel samples used as human productions.

	/i/	/ɪ/	/ɛ/	/ʌ/	/ɑ/
Mean Rating	6.97	4.67	5.89	5.39	5.83
Fundamental Frequency	124.5	132.0	125.0	125.0	125.0

until a frequency beat was observed auditorily. Further adjustments were made until the oscilloscope display showed beats of less than 1 cps. The frequency of the oscillator was then read from a Hewlett Packard 5212-A electronic counter and taken as the fundamental frequency of the vowel sample. Each vowel was measured twice. The mean of these two readings was compared to 130 cps, and if it was within ± 1 semitone, the sample was used as the human production of the vowel. The sample of the vowel /ɪ/ with the highest mean score failed to meet this criterion. Accordingly, the sample with the next highest score was measured. This sample met the criterion and was used for that vowel. In all other cases the sample with the highest mean score was acceptable in terms of fundamental frequency. The obtained frequencies for the five vowels selected are also found in Table 1.

Synthetic vowels.--Two-formant productions of the five vowels /i/, /ɪ/, /ɛ/, /ʌ/, and /ɑ/ and the four intermediate vowels (X_1--X_4) were synthesized. Only two formants were used in order to obtain vowels which could be described in a simple manner with respect to their acoustic characteristics. DeLattre, et al. (1952) and Miller (1953) have reported that two formant vowels can be readily and reliably recognized as the intended vowels. The formant frequencies and their associated bandwidths are seen in Table 2. As described previously, the formant frequencies of the five vowels /i/, /ɪ/, /ɛ/, /ʌ/, and /ɑ/ were taken directly from the averages of the preferred samples of Fairbanks and Grubb (1961). Those for the four intermediate vowels were derived from these data by the sectioning technique. The bandwidths represent the average of the

Table 2. Formant frequencies, bandwidths, and F_1/F_2 ratios for nine synthetic stimuli used as initial and final vowels.

		/i/	/ɪ/	X_1	X_2	/ɛ/	X_3	X_4	/ʌ/	/ɑ/
First Formant	Center Frequency in cps	260	380	410	450	490	520	550	590	780
	Bandwidths in cps	45	45	45	45	45	45	45	45	45
Second Formant	Center Frequency in cps	2 380	2 040	1 910	1 780	1 660	1 490	1 340	1 200	1 060
	Bandwidths in cps	70	70	70	70	70	70	70	70	70
F_1/F_2 Ratio		.109	.186	.215	.253	.295	.349	.410	.492	.736

values reported by Dunn (1961) for these five vowels.

The synthetic vowels were produced by the Communication Sciences Laboratory vowel synthesizer. This device consists of a voice source (a transistorized asymmetrical square-wave generator) driving two cascaded L-C resonant circuits with interstage isolation.

Decade capacitors in the resonant circuits allow the peak frequencies of each of the formant sections to be varied. Bandwidths are similarly adjustable by means of variable resistances in each circuit. The transfer function of the filter section was adjusted with the aid of a Brüel and Kjear 1014 beat frequency oscillator, a Brüel and Kjear 2112 audio frequency spectrometer, and a Hewlett Packard 5212A electronic counter.

In order to obtain the desired stimulus duration the output of the above system was controlled in the same manner as was described for the human vowels. In brief, system on-time was 318 milliseconds ± 1 millisecond; the rise and decay times, 25 milliseconds. The vowel segments were recorded on a Magnecord M-90 tape recorder.

In order to demonstrate that the synthetic productions of the five lead vowels could be expected to be identified as the intended vowels, ten practice items and twenty-five test items were presented to the judges used in the human vowel selection procedure. They were asked to indicate, by a forced-choice procedure, whether each sample was an /i /, /ɪ /, /ɛ /, /ʌ /, or /ɑ /. The results indicated that all of the productions were readily identified. Three of the vowels /i /, /ʌ /, and /ɑ / were identified correctly 100 per cent of the time; intelligibility scores for the /ɪ / and /ɛ / were 96 per cent and 92 per cent, respectively. The four intermediates, X_1--X_4, were

not included in this procedure since they were to be presented in a counterbalanced order and the responses to each would be evaluated in terms of all other items.

Inter-vowel periods

One factor which might be expected to effect the strength of the perceptual shift of a vowel due to an adjacent vowel is the period separating the two. While it was beyond the scope of this investigation to evaluate this temporal effect in detail, a rough attempt was made to obtain evidence regarding the existence of such variation. To this end, each possible pairing of initial and final vowel items were presented with each of two inter-vowel periods, .1 and .5 seconds. The .1 second interval was near the lower limit which could be used with the tape editing technique utilized in the research. The .5 second interval was a convenient multiple of the shorter interval, such a five-fold increase was judged to be sufficient to provide a reasonably adequate difference in the sampling points.

Preparation of experimental tapes

The experimental tape was constructed by splicing together in random order two samples of each possible combination of initial vowel source, initial vowel, inter-vowel period, and final vowel. The pairs were constructed in the following manner. The initiation and termination of a sample were located utilizing a Minnesota Mining and Manufacturing Company tape viewer. For an initial vowel, the tape was marked at a point approximately one-half second before the initiation of the signal and as carefully as possible at a point equivalent to either .05 or .25 seconds (depending on the temporal condition) after the termination. The following vowel for that item

was marked at a point preceding the vowel by an amount equal to one half the inter-stimulus period and at a point approximately one-half second after the termination of the vowel. The two vowels used were spliced together at these points to form an item pair. In turn, the pairs were spliced to leader which formed the judging intervals. Items number 121 through 140 were duplicated and used for practice items.

Listener selection

The twenty-five listeners used in the study were members of the faculty or students at the Communication Sciences Laboratory and the Speech Department at the University of Florida. All were speakers of American English and exhibited essentially normal hearing. As a single exception, one subject exhibited a monaural high-frequency loss; however, his loss was above the range usually considered important in the perception of speech. All listeners were skilled in the use of the International Phonetic Alphabet.

Potential subjects were screened for ability to perform the task. A subject screening tape consisting of twenty practice items and seventy-five test items taken from the vowel selection tape was presented to all potential subjects.[1] All listeners selected correctly identified at least 90 per cent of the test items.

Experimental procedure

The twenty-five listeners were seated in the IAC room described previously, in groups of one to four, and the "Instructions to

---

[1]See Appendix A for Listener's Screening Form.

Listeners" (Appendix B) were read. Briefly, they were instructed to attend to each vowel pair and decide which of the five vowel categories / i /, / ɪ /, / ɛ /, / ʌ /, or / ɑ / best described the second vowel in the set. Each listener received a response form and a marking template.

The experimental tape was played on a Magnecorder M-90 tape recorder coupled to a Marantz Model 7 preamplifier and 8B power amplifier driving an Acoustic Research AR-3 speaker system. The twenty practice items were run and, after a brief interval during which questions were answered, the experimental tapes were presented.

Data reduction and analysis

The test forms were scored and distributions of responses tallied for each item by an IBM Model 1230 test scoring machine and an IBM Model 1401 computer. Statistical analyses were carried out on an IBM Model 709 computer.

In order to be able to consider the obtained data statistically, the responses / i /, / ɪ /, / ɛ /, / ʌ /, and / ɑ / were assigned numbers from one through five, respectively, and treated as ordinal quantities. This approach was judged to be justified for two reasons. First, the vowels fall in this specific order on the two acoustic continua which are judged to be most significant in vowel quality, i.e., formants one and two. Second, the responses obtained in this study demonstrate that, with scattered exceptions, the "error" or non-majority responses were in categories adjacent to the "correct" category as would be expected if the assumption of ordinality were justified.

In order to evaluate inter- and intra-observer reliability Spearman rank order correlations were calculated; (1) between replications for each listener and (2) between listeners. The intra-judge correlations ranged from .61 to .96 with a mean of .88. The distribution was markedly skewed due to one very low score (.61). On the basis of the inter-correlations it seems probable that some external factor unduly influenced this judge's responses to a large number of items in the early portion of the experimental procedure. However, in the absence of external evidence of this effect, these data could not be removed from the statistical analysis.

Inter-judge reliability ranged from .51 to .97 with a mean of .85. The first replication for the same judge is again responsible for the lower scores. In sum, however, it was judged that these data indicate the overall reliability of the judges to be within acceptable limits.

In order to be able to test the significance of observed shifts statistically with a chi square procedure, the responses were also scored as a trivariate scheme (plus, equal, or minus). A plus was used to indicate that the F_1/F_2 ratio of the stimuli being judged was higher than that of the vowel used to describe it. A zero, that the response was "correct," and a minus, that the F_1/F_2 of the judged stimuli was lower than that of the vowel used to describe it. Statistical tests for the significance of the various effects were performed and the data were presented graphically to display the relative magnitudes and directions of the observed changes.

RESULTS

Twenty-five listeners were asked to respond to 280 vowel pairs. Each possible combination of human and synthetic initial vowel source, five initial vowels, two inter-vowel periods, and the seven final vowels, was presented twice in randomized order. Listeners were required to identify the final vowel of each pair in a closed set of five vowels.

Table 3 summarizes the subjects' responses to each pairing of initial and final vowels.[1] As would be expected the two-formant stimuli which were generated to represent the /ɪ/, /ɛ/, and /ʌ/ vowels were most often identified appropriately by listeners. Moreover, stimuli with intermediate formant characteristics (X_1--X_4) were identified as having ambiguous auditory characteristics. From the proportions of responses in the two adjacent categories, it would seem that X_2 and X_3 almost perfectly bisected their respective adjacent vowels while X_1 and X_4 apparently were better examples than intended of the nearest vowel.

Additionally, it may be seen that, with few exceptions, all of the responses for each following vowel fell in either one or two main response categories with only an occasional scattered response in a

[1]Scores for the two initial vowel sources (human and synthetic) and two inter-vowel periods (.1 and .5 seconds) have been pooled in this table. Statistical evidence justifying this pooling is presented in a later section.

Table 3. Numbers of responses in each of five categories for each initial vowel, following vowel combination.

Following Vowel	Initial Vowel	Response				
		/i/	/ɪ/	/ε/	/ʌ/	/ɑ/
/ɪ/	/i/	3	190	6		1
	/ɪ/	20	173	3	3	1
	/ε/	32	166	2		
	/ʌ/	29	169	1	1	
	/ɑ/	26	171	1	2	
X_1	/i/	1	185	14		
	/ɪ/	3	179	14	4	
	/ε/	9	188	2	1	
	/ʌ/	10	187	2	1	
	/ɑ/	4	187	9		
X_2	/i/	1	90	106	2	1
	/ɪ/		60	137	3	
	/ε/		163	32	5	
	/ʌ/	1	134	60	5	
	/ɑ/		115	80	5	
/ε/	/i/		9	165	25	1
	/ɪ/		9	177	10	4
	/ε/	1	16	163	17	3
	/ʌ/		12	171	16	1
	/ɑ/		17	158	24	1
X_3	/i/		5	97	96	2
	/ɪ/		1	71	122	6
	/ε/		6	97	95	2
	/ʌ/		3	115	81	1
	/ɑ/		4	111	85	
X_4	/i/		1	17	180	2
	/ɪ/		2	11	182	5
	/ε/			11	187	2
	/ʌ/			16	178	6
	/ɑ/			21	178	1
/ʌ/	/i/		1		179	20
	/ɪ/		1		173	26
	/ε/			2	163	35
	/ʌ/	1	1	1	170	27
	/ɑ/				191	9

cell not immediately adjacent to the main cell(s). In other words, errors tended to be distributed about the modal response in the pattern which would be expected for responses to ordinal stimuli.

It has been hypothesized that the effect of a vowel on the perception of a following vowel is one of contrast. That is, the difference between two vowels is perceived as greater than it actually is with respect to the F_1/F_2 continuum. In order to test the hypothesis that the contextual effect of one vowel upon another actually was one of contrast, the obtained responses were recast into a two-way classification in the following manner. If the F_1/F_2 ratio[1] of the vowel used to describe a stimulus was greater than the specified F_1/F_2 ratio for that stimulus, the response was categorized as "plus" and placed in the "greater" category. The responses in which the F_1/F_2 ratio of the response vowel was equal to that of the sample were categorized "zero" and one half of these values were added to the "greater" cateogry.[2] These data then were transformed to proportions for statistical analysis. Specifically, this proportion was defined as $\frac{P + \frac{E}{2}}{N}$ X 100 where P equals the number of "plus" responses, N equals the total number of responses, and E equals the number of zero (or equal) responses. For the / i /--/ ɪ /

---

[1]See Table 2 again for listings of the F_1/F_2 ratios.

[2]Naturally the residual would represent "lesser" responses but since these data are the single inverse of the "greater" responses, they would show identical types of patterns. Accordingly, all consideration of results is based on "greater" responses only.

pairs this value was $\frac{3 + \frac{190}{2}}{200}$ X 100 = 49.00. Since there could be no equal responses for the four intermediate vowels the above formula was simplified to $\frac{P}{N}$ X 100 for these events. Accordingly, the proportion of "greater" responses for the / i /--X_1 pairs was $\frac{186}{200}$ X 100 = 93.00. The proportions of greater responses for each initial vowel-following vowel combination are seen in Table 4. It would be predicted, from the hypothesis of a contrast effect, that the proportion of greater responses would vary as a function of the relative F_1/F_2 ratios of the two vowels in a pair. Specifically, if the F_1/F_2 ratio of the initial vowel was higher than that of the following vowel, the proportion of greater responses should be reduced while for the opposite case the proportions should be larger. The relative sizes of these proportions for each following vowel do vary as a function of the initial vowel. For example, when / ɪ / was preceded by / i /, the proportion of greater responses was less than for any other initial vowel. For following vowel X_1, the proportions with both / i / and / ɪ / as initial vowels were less than those for the remaining three cells. This pattern is found to repeat for each following vowel, the dividing line between high and low cells shifting as a function of the following vowel. This is more easily seen when these data are presented graphically in Figure 3a-g. These figures show the proportions of greater responses for each final vowel as a function of initial vowel. A vertical line has been drawn through each figure at a point appropriate to the ordinal position of the following vowel (in relation to the initial vowel), thus the graph is divided into two markedly different segments. While there is considerable variation in the magnitude of the

Table 4. Proportions of greater responses for each initial vowel-following vowel combination.

Following Vowel	Initial Vowel	Proportion of Greater Responses
/ɪ/	/i/	49.00
	/ɪ/	53.25
	/ɛ/	57.50
	/ʌ/	56.75
	/ɑ/	55.75
X_1	/i/	93.00
	/ɪ/	91.00
	/ɛ/	98.50
	/ʌ/	98.50
	/ɑ/	95.50
X_2	/i/	45.50
	/ɪ/	30.00
	/ɛ/	81.50
	/ʌ/	67.50
	/ɑ/	57.50
/ɛ/	/i/	45.75
	/ɪ/	48.75
	/ɛ/	49.25
	/ʌ/	48.75
	/ɑ/	48.00
X_3	/i/	51.00
	/ɪ/	36.00
	/ɛ/	51.50
	/ʌ/	59.00
	/ɑ/	57.50
X_4	/i/	9.00
	/ɪ/	6.50
	/ɛ/	5.50
	/ʌ/	8.00
	/ɑ/	10.50
/ʌ/	/i/	45.25
	/ɪ/	43.75
	/ɛ/	41.75
	/ʌ/	44.00
	/ɑ/	47.75

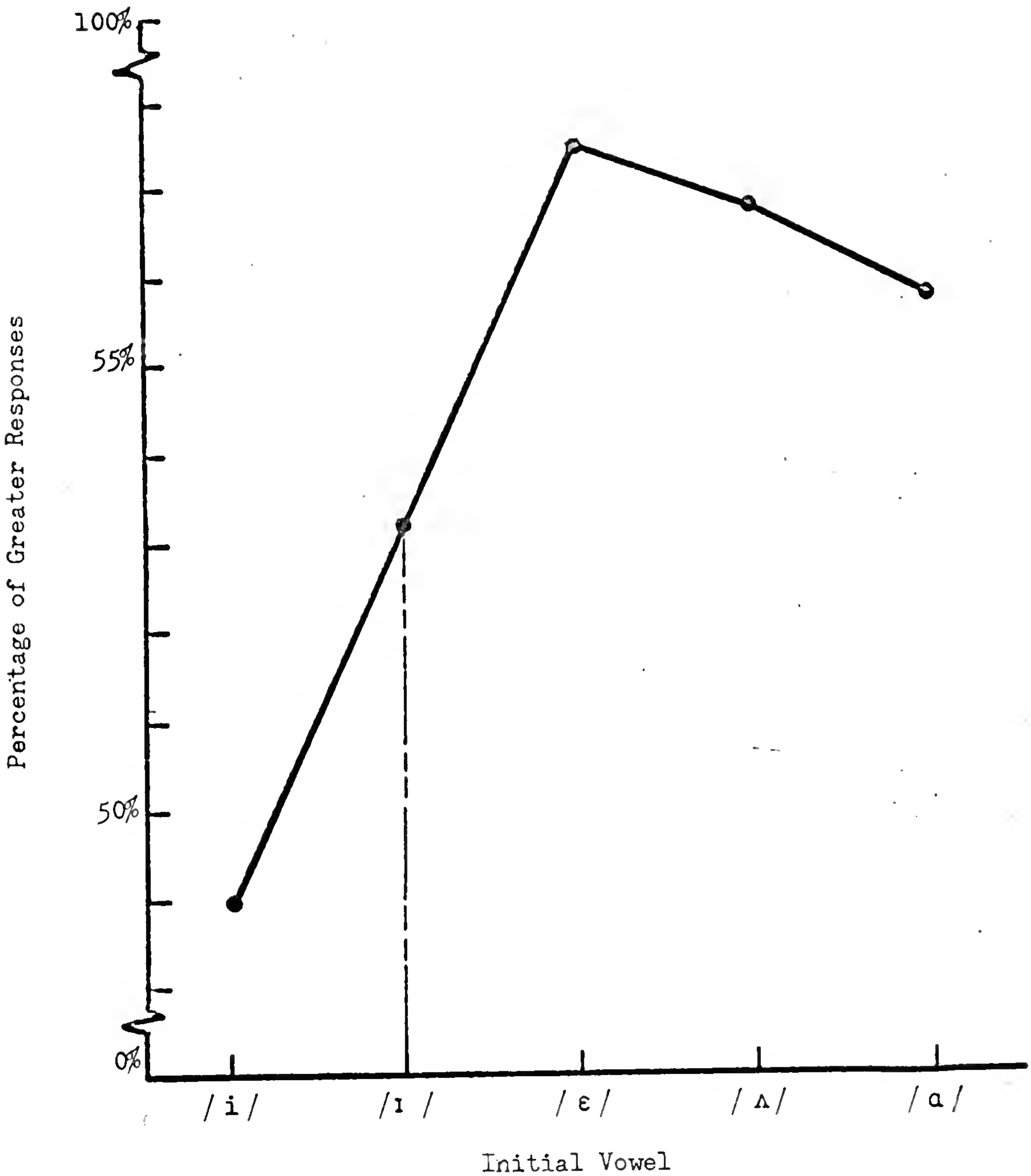

Figure 3a. Percentage of greater responses as a function of the initial vowel for the following vowel / ɪ /.

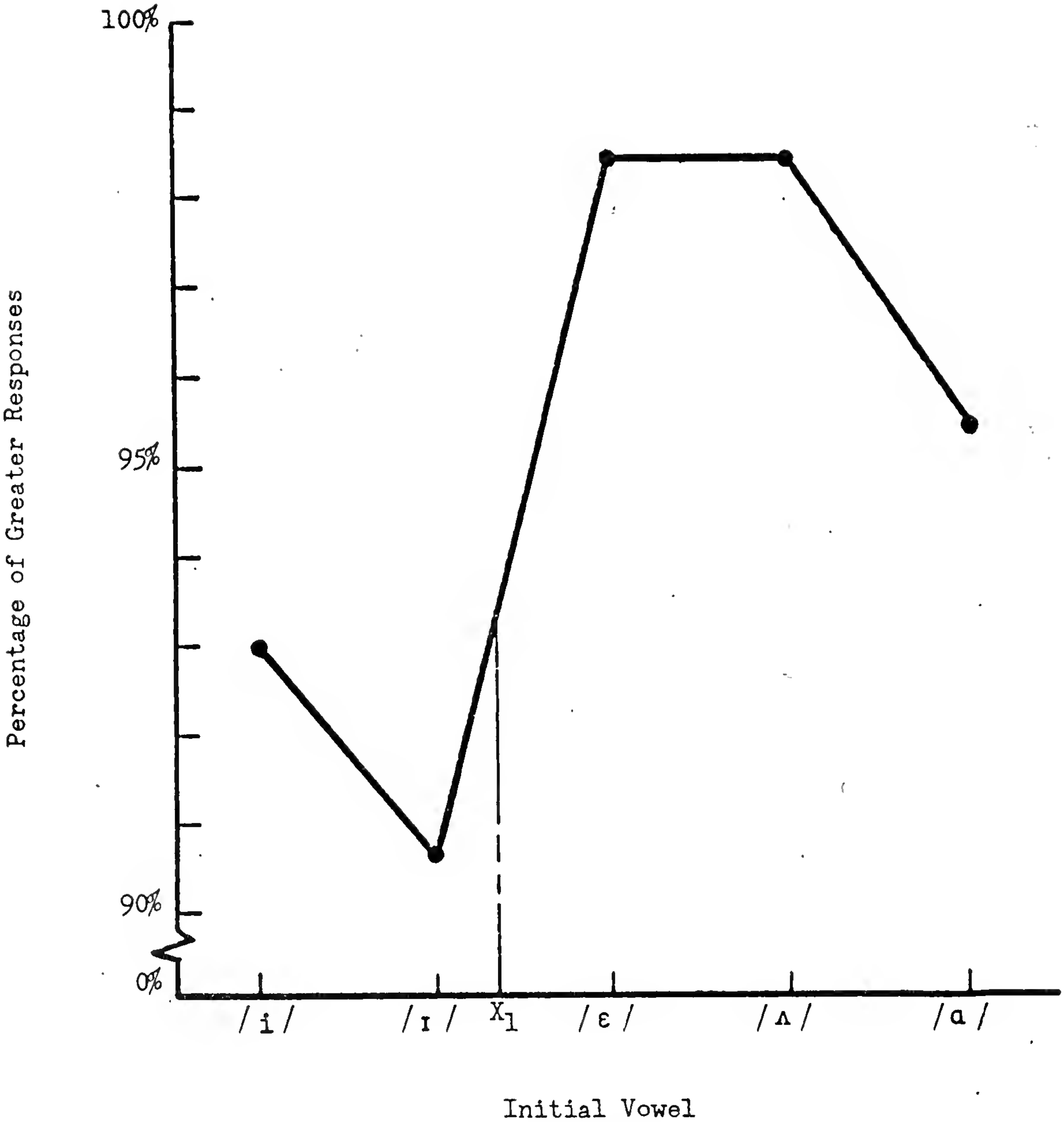

Figure 3b. Percentage of greater responses as a function of the initial vowel for the following vowel X_1.

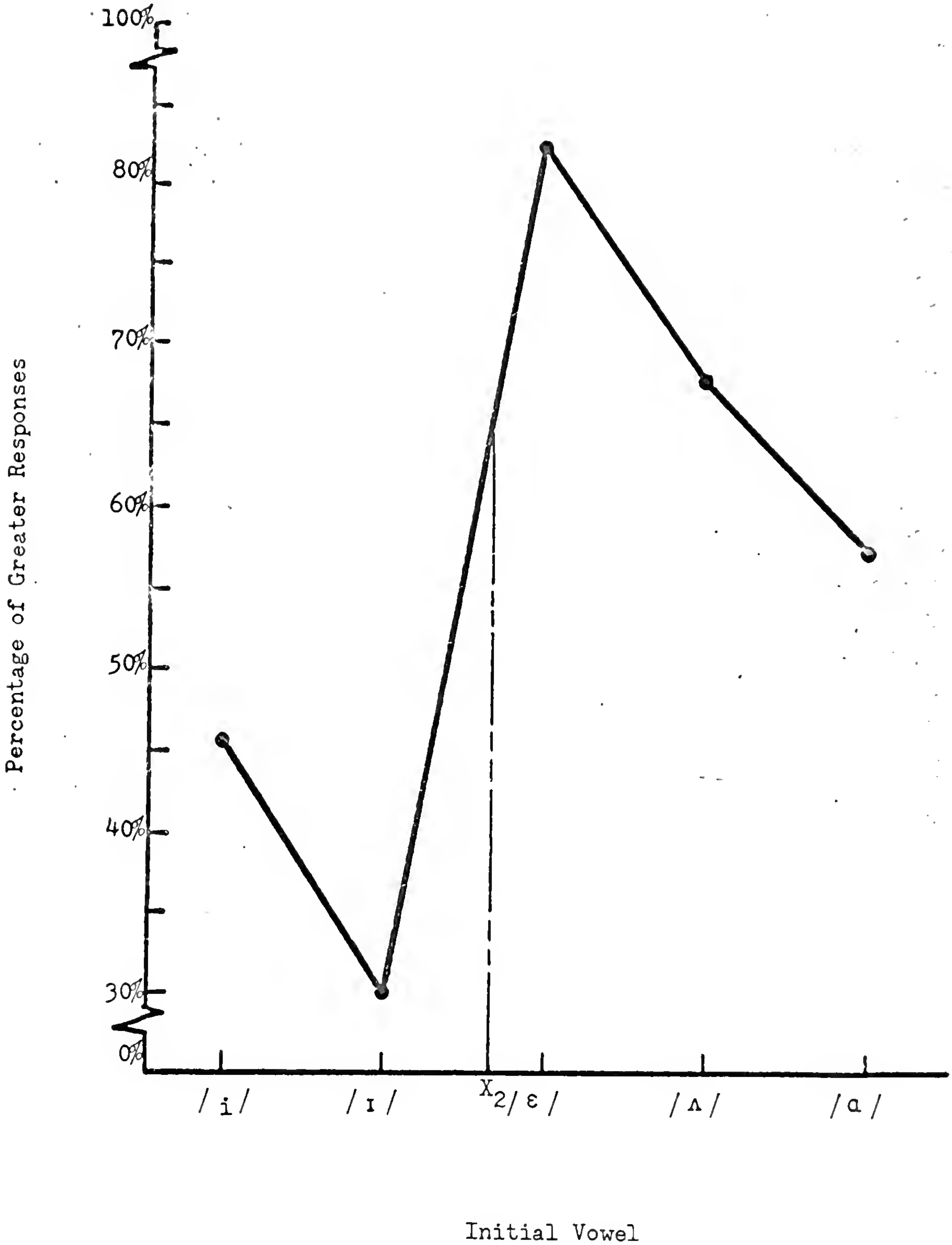

Figure 3c. Percentage of greater responses as a function of the initial vowel for the following vowel X_2.

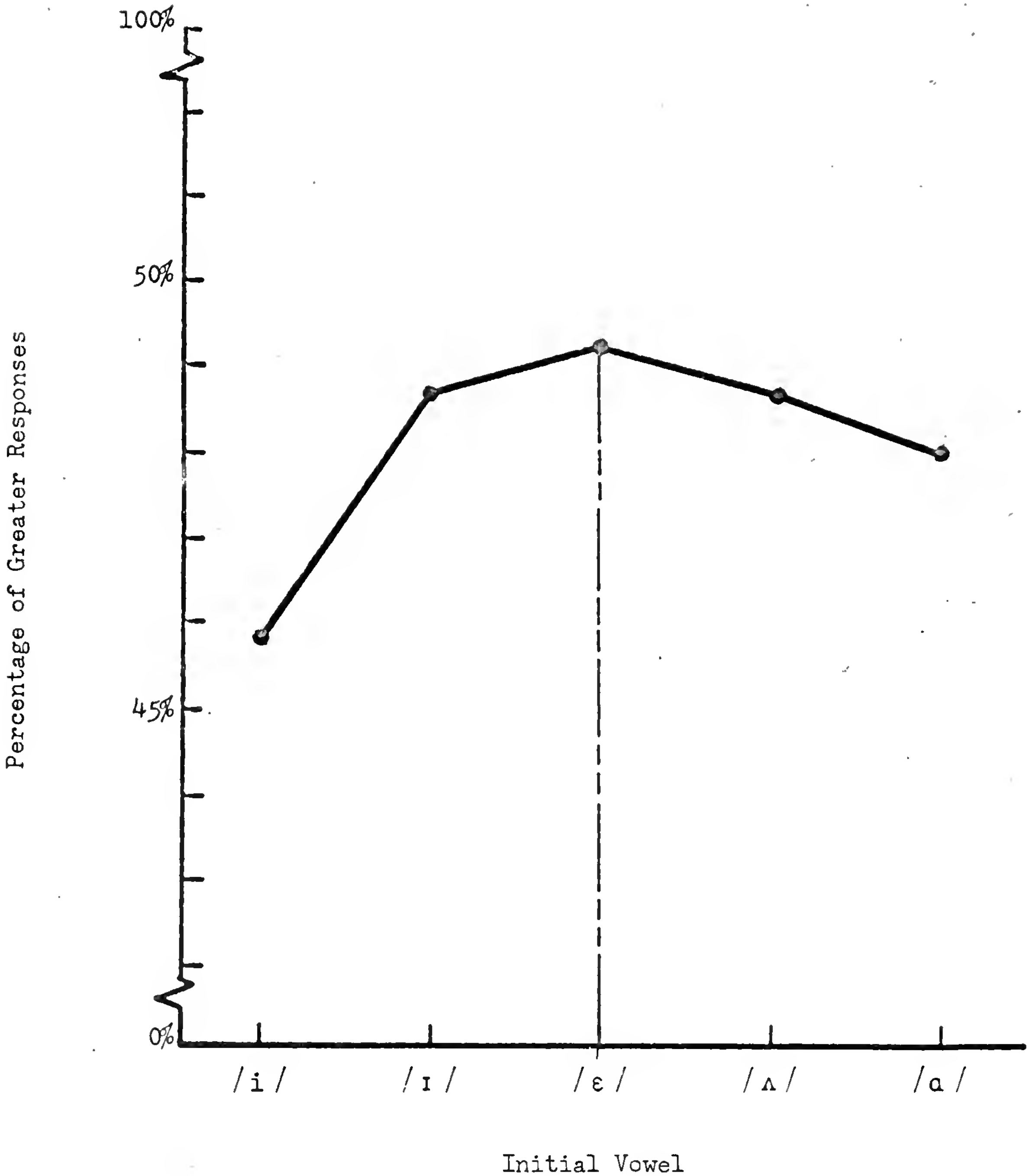

Figure 3d. Percentage of greater responses as a function of the initial vowel for the following vowel / ɛ /.

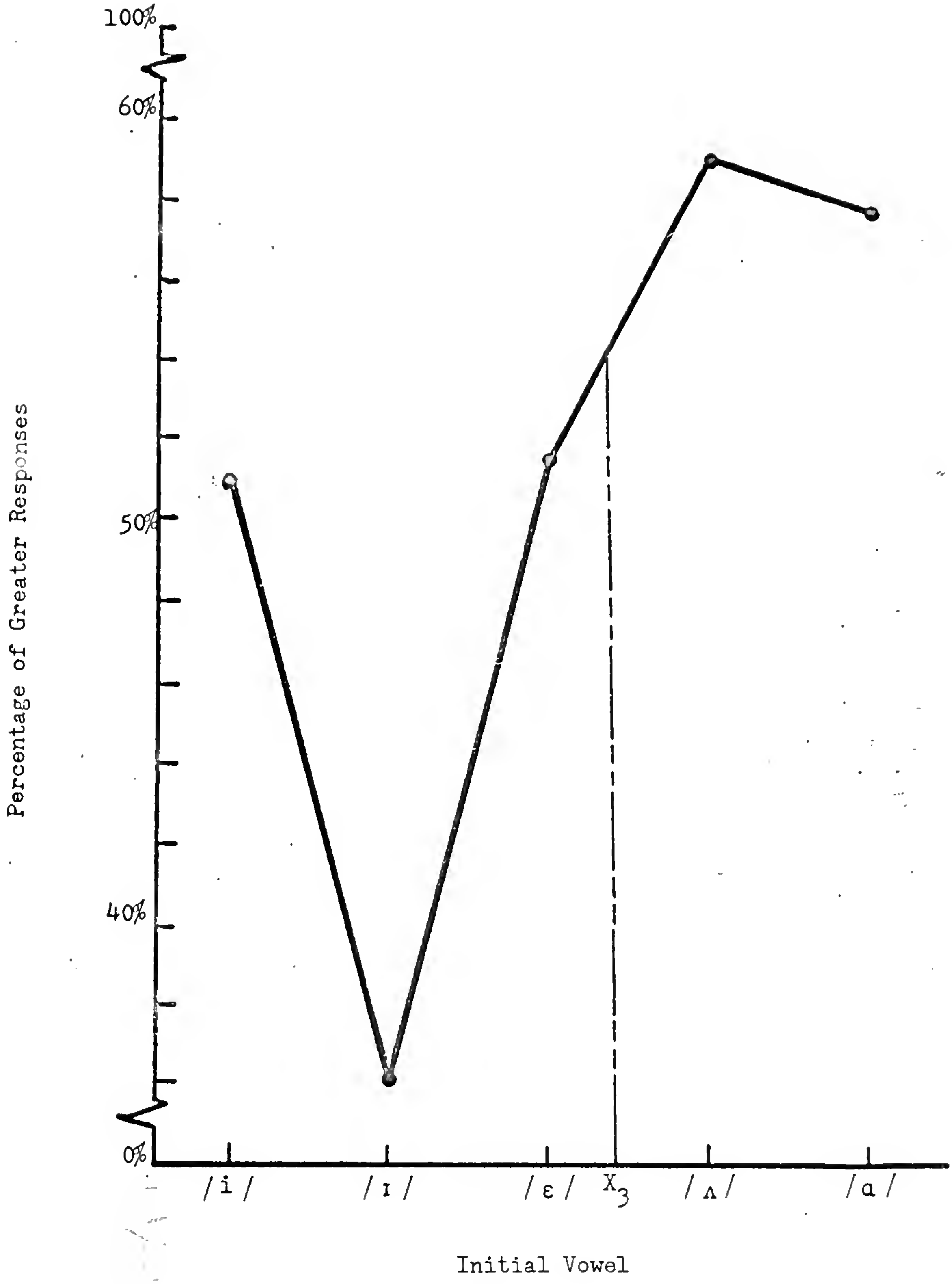

Figure 3e. Percentage of greater responses as a function of the initial vowel for the following vowel X_3.

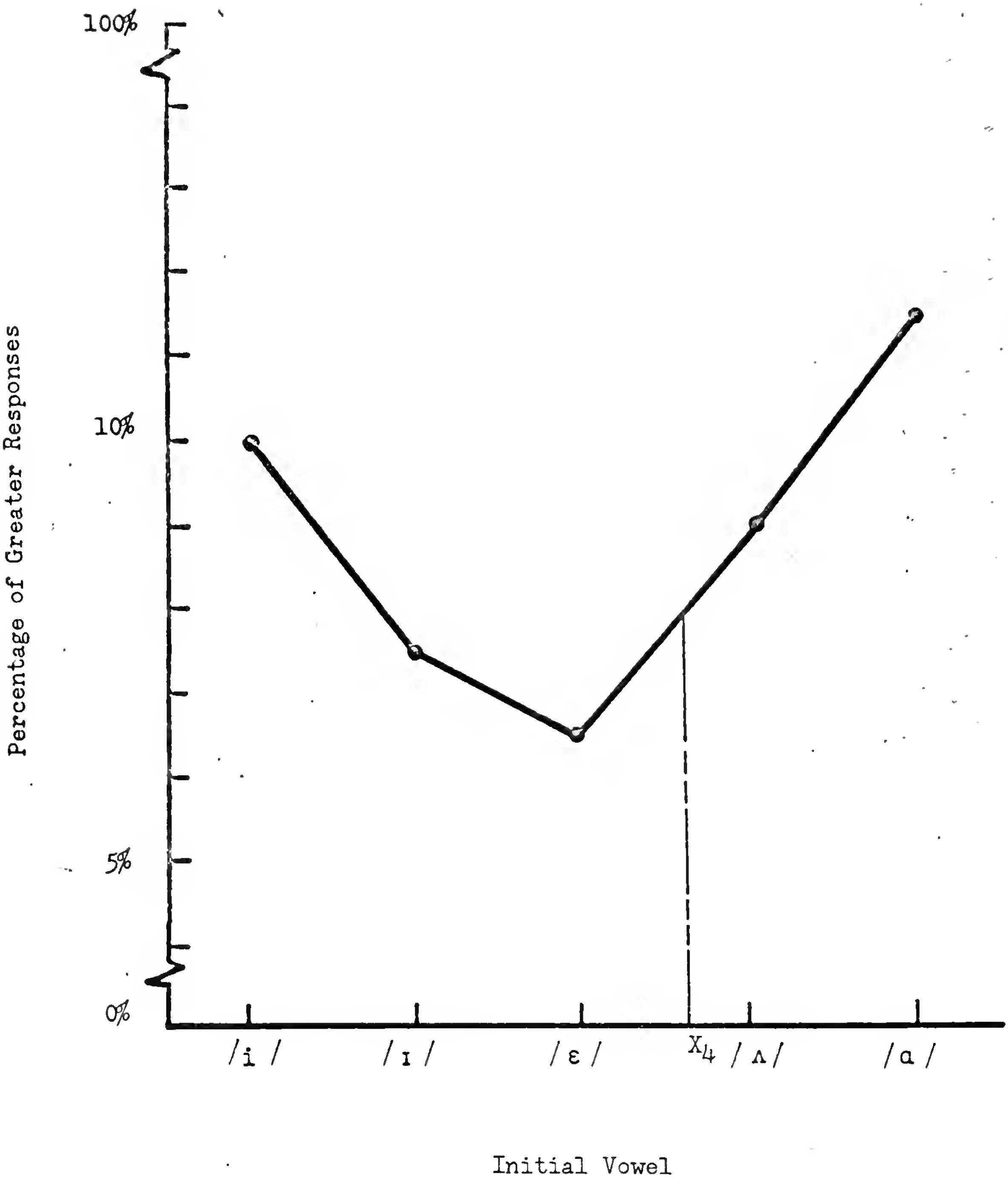

Figure 3f. Percentage of greater responses as a function of the initial vowel for the following vowel X_4.

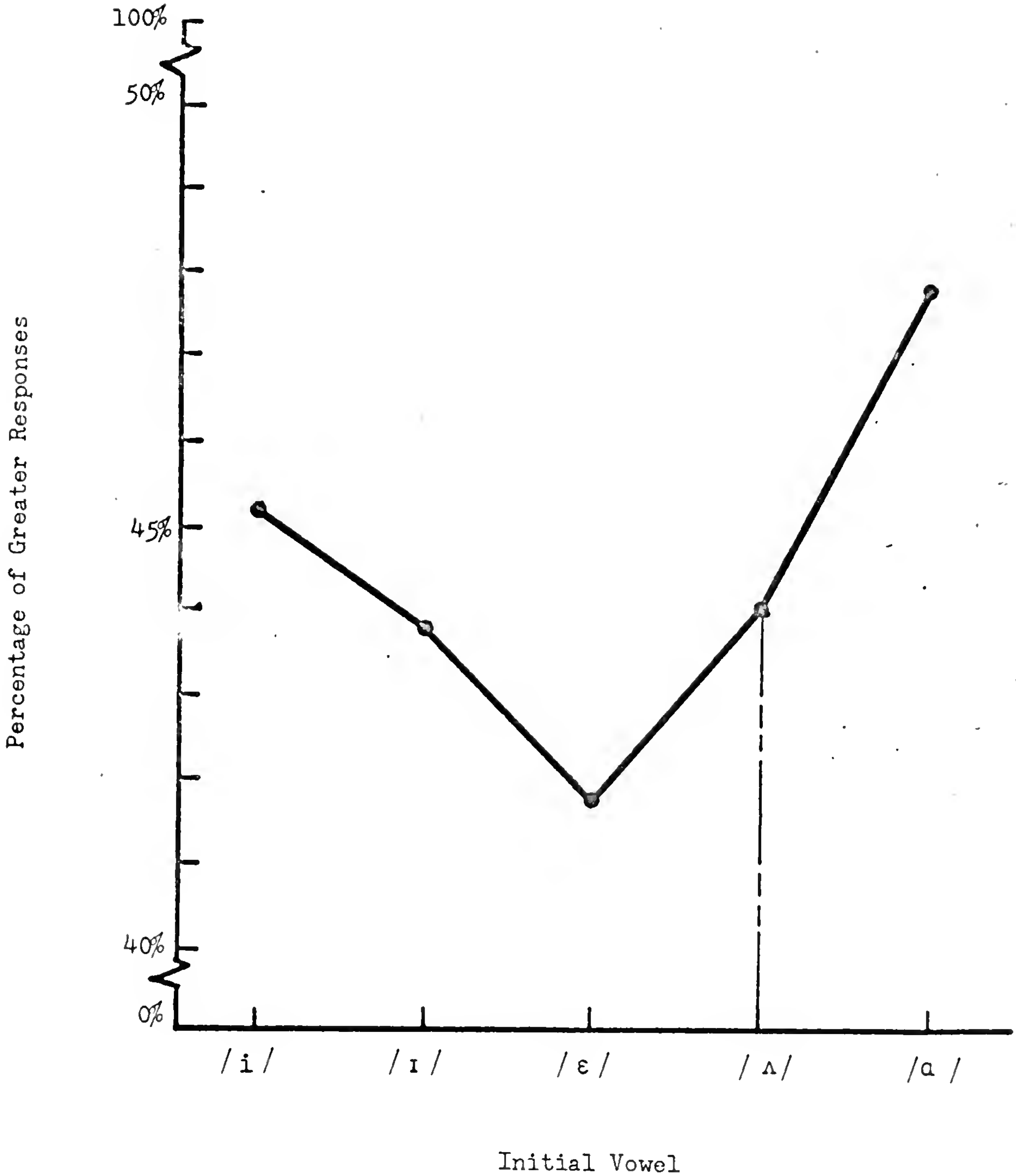

Figure 3g. Percentage of greater responses as a function of the initial vowel for the following vowel /ʌ/.

difference for the seven following vowels, an overall pattern is readily seen. With the exception of the results for the following vowel /ɛ/, the data points are almost unanimously as predicted, i.e., lower when the F_1/F_2 ratio of the initial vowel is higher than that of the following vowel, and higher when the opposite relationship exists. The consistency with which the predicted pattern is observed undoubtedly indicates that the contextual effect of vowel on vowel--at least for these phonemes--is one of contrast. Further, it seems that the effect varies considerably as a function of the ambiguity of the following vowel, being greatest when it is relatively ambiguous. Finally, there is also a definite tendency for the greatest shifts to occur when the initial and final vowels are relatively similar. However, this finding was not without exception (for example, the /ɛ/--X_3 and /ʌ/--X_4 pairs).

In Figure 4 the data for all seven following vowels have been summarized. The pairs were divided into seven categories, -3 to +3 with "equal" responses scored a zero. The matrix used to categorize the vowel pairs into these sets is seen in Table 5. In order to obtain data for use in this figure, the proportion for each pairing was subtracted from the mean proportion for its following vowel. For example, the mean proportion for the following vowel X_2, was 56.40. Thus, the values entered in the matrix were for -2, -10.90; for -1, -26.40; for +1, 25.10; for +2, 11.10; and +3, 1.10. These represent the differences from the mean for the pairing of X_2 with each of the five initial vowels. Since the number of entries for the various vowel pair difference categories varied, the sum for each column was divided by the number of entries in that column to obtain the values plotted in Figure 4.

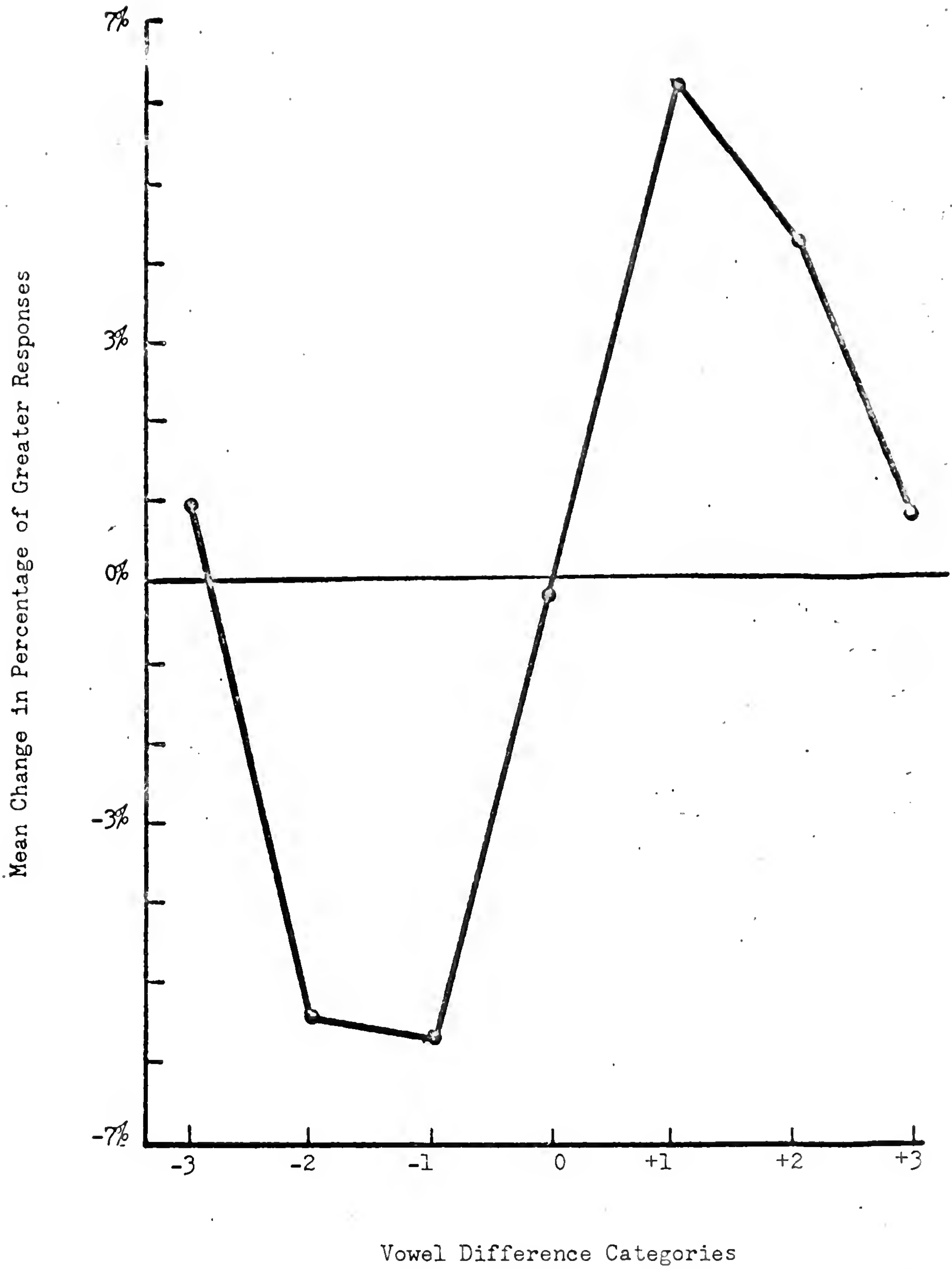

Figure 4. Summary of data for all vowel pairs showing the mean change in percentage of greater responses for seven categories of initial vowel-following vowel difference.

Table 5. Vowel pair difference categories for the five initial vowels with each following vowel.

Following Vowel	Vowel Pair Difference Category						
	-3	-2	-1	0	+1	+2	+3
/ɪ/			/i/	/ɪ/	/ɛ/	/ʌ/	/ɑ/
X_1		/i/	/ɪ/		/ɛ/	/ʌ/	/ɑ/
X_2		/i/	/ɪ/		/ɛ/	/ʌ/	/ɑ/
/ɛ/		/i/	/ɪ/	/ɛ/	/ʌ/	/ɑ/	
X_3	/i/	/ɪ/	/ɛ/		/ʌ/	/ɑ/	
X_4	/i/	/ɪ/	/ɛ/		/ʌ/	/ɑ/	
/ʌ/	/i/	/ɪ/	/ɛ/	/ʌ/	/ɑ/		

The pattern of the contrast effect is seen more clearly in this figure. Specifically, the effect of an initial vowel upon the identification of the following vowel is seen to be greatest for samples which were closely adjacent with respect to F_1/F_2 ratios and tends to decay as the following vowel becomes less like the initial vowel. This relationship constitutes the basic finding of this research. However, as was noted, the magnitude of this contextual effect is not great. For example, even though the range for the ambiguous vowel X_2 was from 30-81 per cent, Figure 4 demonstrates that the total extent of the average proportional shift was only 12 per cent.

While the above mentioned effect of an initial vowel upon the identification of a following vowel has been demonstrated by the data presented, it was desirable that both this effect and the effects of inter-vowel interval and initial vowel source be subjected to statistical confirmation. Accordingly, an analysis of variance was carried out. It must be realized that this statistical procedure was undertaken, even though two of the basic assumptions were violated. That is, neither the assumption of interval data nor that of homogeneity of variance, both of which are essential to the proper use of an analysis of variance, was justified. First, it can not be assumed that the five response levels fall at equal intervals along a continuum scale. Second, a Bartlett's Test for homogeneity of variance within the 280 test items gave a chi square value of 337.5 for 266 degrees of freedom. The probability of the occurrence of such a high value, if the variance within the cells was homogeneous, is very low ($p < .001$). Nevertheless, use of an analysis of variance

procedure seems justified as long as it is realized that probability statements are approximate and may be somewhat inflated.

Table 6 presents the results of an analysis of variance used to evaluate the four main effects (initial vowel source, initial vowel, inter-vowel interval, and following vowel) and their interactions. As expected, the F for the following vowel is extremely large. As this is the vowel that the subjects were instructed to classify, it would have been suprising had this test not reached significance. The next largest F ratio is for initial vowel. This value ($F = 24.87$, $df = 4$ and 24) substantially exceeds that necessary to achieve a .001 confidence level. This confirms the statistical significance of the shifting related to following vowel identification as a function of differing initial vowels. The interaction of initial and final vowels also resulted in an F ratio which exceeded the value needed for significance at the .001 level: a finding which also tends to confirm the statistical significance of the observed shift. Both of these relationships are logically attributable to the differential effects of the initial vowels on the identification of the following vowels--due to their relative F_1/F_2 ratios.

The main effect of the initial vowel would be predicted because both /i/ and /ɪ/ could only act to increase the mean responses to the various following vowels while /ʌ/ and /ɑ/ could only lower the mean scores. The effect of /ɛ/ would be neutral, that is, it would tend to cancel itself out by increasing the means for some of the vowels and decreasing it for others. Similarly, the interaction of initial and final vowels would be predicted from the hypothesized contrast effect. Specifically, (1) the variation of the magnitude of

Table 6. Summary of an analysis of variance for the main effects and interactions of initial vowel source, initial vowel, inter-vowel period, and following vowel.

Source	df	M. S.	F
S	1	.002	.013
I	4	4.391	24.87***
P	1	.514	2.91
F	6	786.394	4 455.49***
S X I	4	.724	4.10**
S X P	1	.586	3.32
S X F	6	.257	1.46
I X P	4	.909	5.15***
I X F	24	1.348	7.64***
P X F	6	.395	2.24*
S X I X P	4	.094	.53
S X I X F	24	.247	1.40
S X P X F	6	.092	.52
I X P X F	24	.226	1.28
S X I X P X F	24	.182	1.03
Replications (including interactions)	140	.237	
Within Cells	6 720	.176	
Total	6 999		

* = P < .05

** = P < .01

*** = P < .001

the effect with the relative difference of the F_1/F_2 ratios of the vowels in a pair and (2) the reversal of its direction (depending upon whether the F_1/F_2 ratio of the initial or final vowel is largest) would result in such an interaction.

While the main effect for period did not reach significance at the $P < .05$ level there was a large interaction for initial vowel and period. Since the contrast effect of two of the initial vowels would be predominately positive, while for two others it would be predominately negative, a tendency for the magnitude of the effect to be consistently greater or smaller at one of the intervals, would result in such an interaction. In order to determine whether or not such a tendency can be inferred from these data, they are presented graphically in Figure 5, as the proportion of "greater" responses for the five initial vowels at each temporal interval. Adjacent to the obtained data are the trends which might have been predicted from the contrast effect. The mean values of the observed and expected curves were equated. The direction of the shift is predicted by the contrast effect and the magnitudes of the changes were arbitrarily selected to conform as closely as possible to the obtained values while showing the relative differences between vowels. Unfortunately, the temporal shift is as predicted for only three of the five vowels. For a fourth, /ɑ/, there is a small reversal from prediction while /ɛ/ shows a marked discrepancy.

These data have been summarized in Figure 6. The matrix of vowel pair difference values, Table 5, was used to categorize the vowel pairs into these seven categories in the same manner as it was used previously. The data plotted represent the means of the

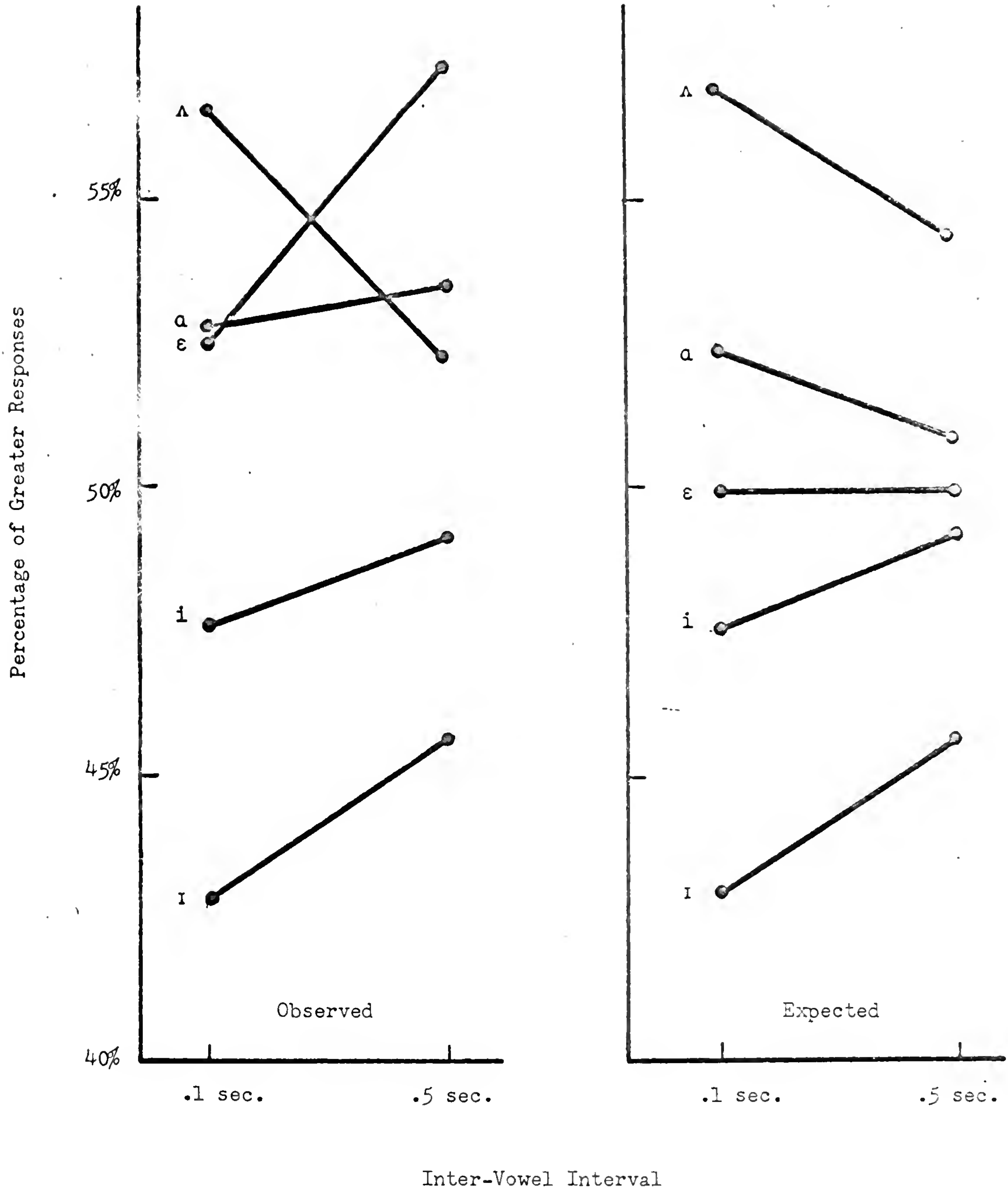

Figure 5. Observed and expected percentages of greater responses for each initial vowel at two inter-vowel intervals.

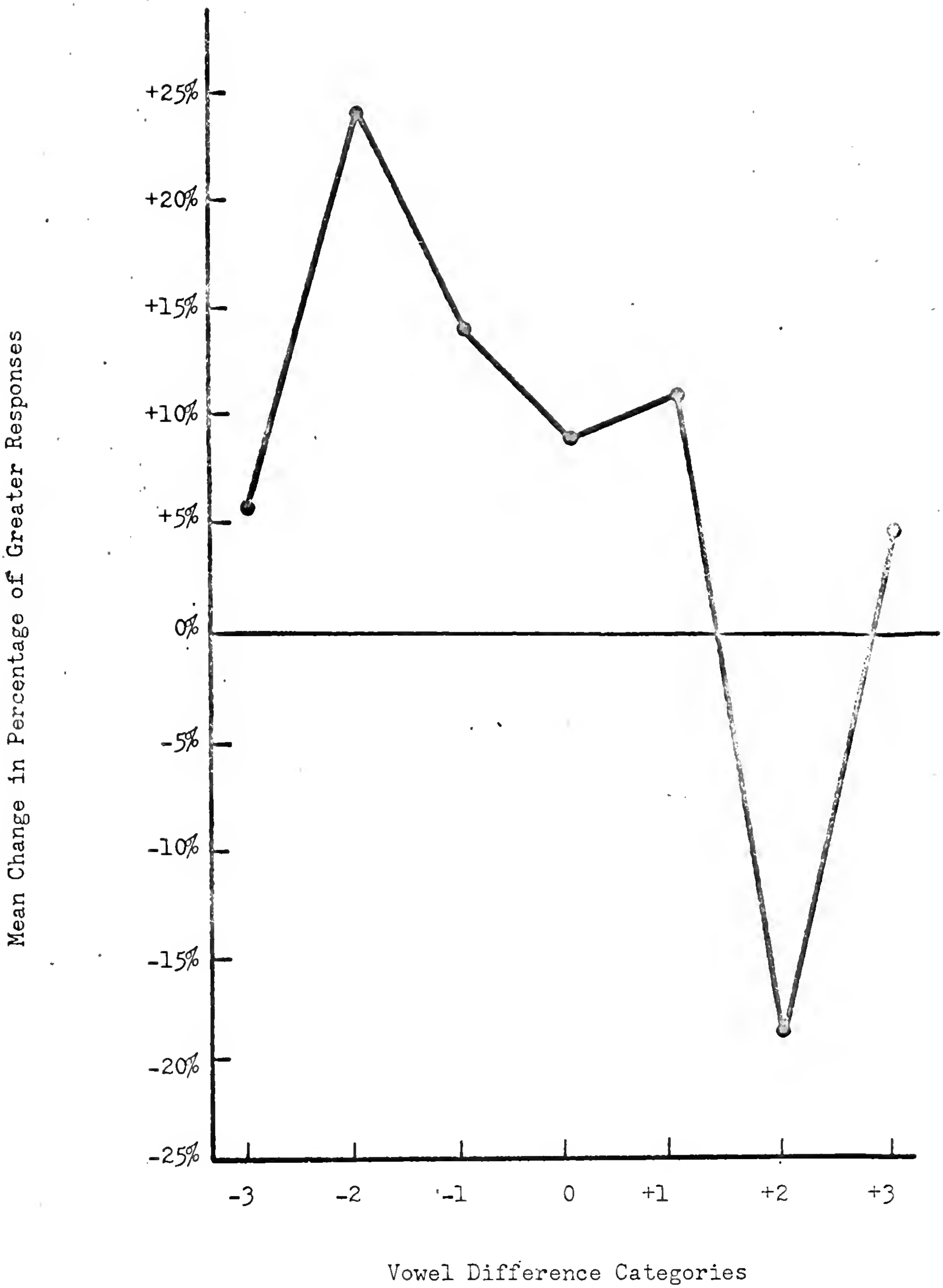

Figure 6. Summary of the data for all vowel pairs showing the mean differences in the percentage of greater responses between two inter-vowel intervals, .1 and .5 seconds, for seven categories of initial vowel-following vowel difference.

differences in proportions of greater responses for the two inter-vowel intervals in each of the seven vowel pair difference categories. A negative value indicates that the proportion of greater responses was larger for the .1 second interval.

From this figure it can be seen that the proportion of greater responses varies as a function of time as would be predicted from the assumption that the contextual effect decreases with an increase in the inter-vowel interval. That is, for the vowel difference categories in which the contextual effect had been negative, the proportions show a regression toward a more neutral value as the interval is increased (the change over time is positive). On the other hand, for categories in which the contextual effect was positive, this pattern is not seen as clearly; however, the tendency is evident with the overall difference in proportions for the two periods being considerably more negative than were those for the cells in which the main contrast effect had been negative.

The F ratio for the interaction between initial vowel and initial vowel source, i.e., human vs. synthetic, was significant at the .01 level. Since the main effect of vowel source was not significant, such a difference is probably explained by the fact that the F_1/F_2 ratios of the human and synthetic productions of each vowel differs slightly. As these differences would not be expected to be from vowel to vowel the overall effect of source would cancel itself out while these minor variations would be seen in the form of this interaction.

The interaction between inter-vowel period and the following vowel reached significance at the .05 level. When the percentage of

greater responses are plotted as in Figure 7 for the seven following vowels as a function of period, it can be seen that these proportions vary for three of the seven vowels, however, no pattern is evident which might relate to the interpretation of the results of the present investigation.

In order to confirm the statistical significance of the major finding of this research, i.e., the shifting of the identification of a vowel away from a preceding vowel, an additional statistical test was performed. Table 7 presents chi square values for the seven following vowels. Each was divided into two cells, one in which the initial vowel F_1/F_2 ratio is greater than that for the following vowel and a second, in which the F_1/F_2 ratio of the following vowel was greater. For the vowels / ɪ /, X_1, X_2, and X_3 the chi square values for these dichotomies were significant at the $P < .001$ level. For the remaining three vowels chi square values were not significant and in fact for two, / ɛ/ and X_4, they were extremely low.

Interpretation of the chi square results would suggest that at least for four of the following vowels the stated dichotomy does define a real difference in the data, while for the remaining three samples the data do not justify such a conclusion. However, examination of Figure 3a-g will reveal the obvious similarity in the patterns of X_4 and / ʌ / to the vowels for which significance was obtained. Such similarity would demonstrate that a similar effect did occur in these vowels even though its lesser strength did not result in differences which were measurable by the statistical test employed.

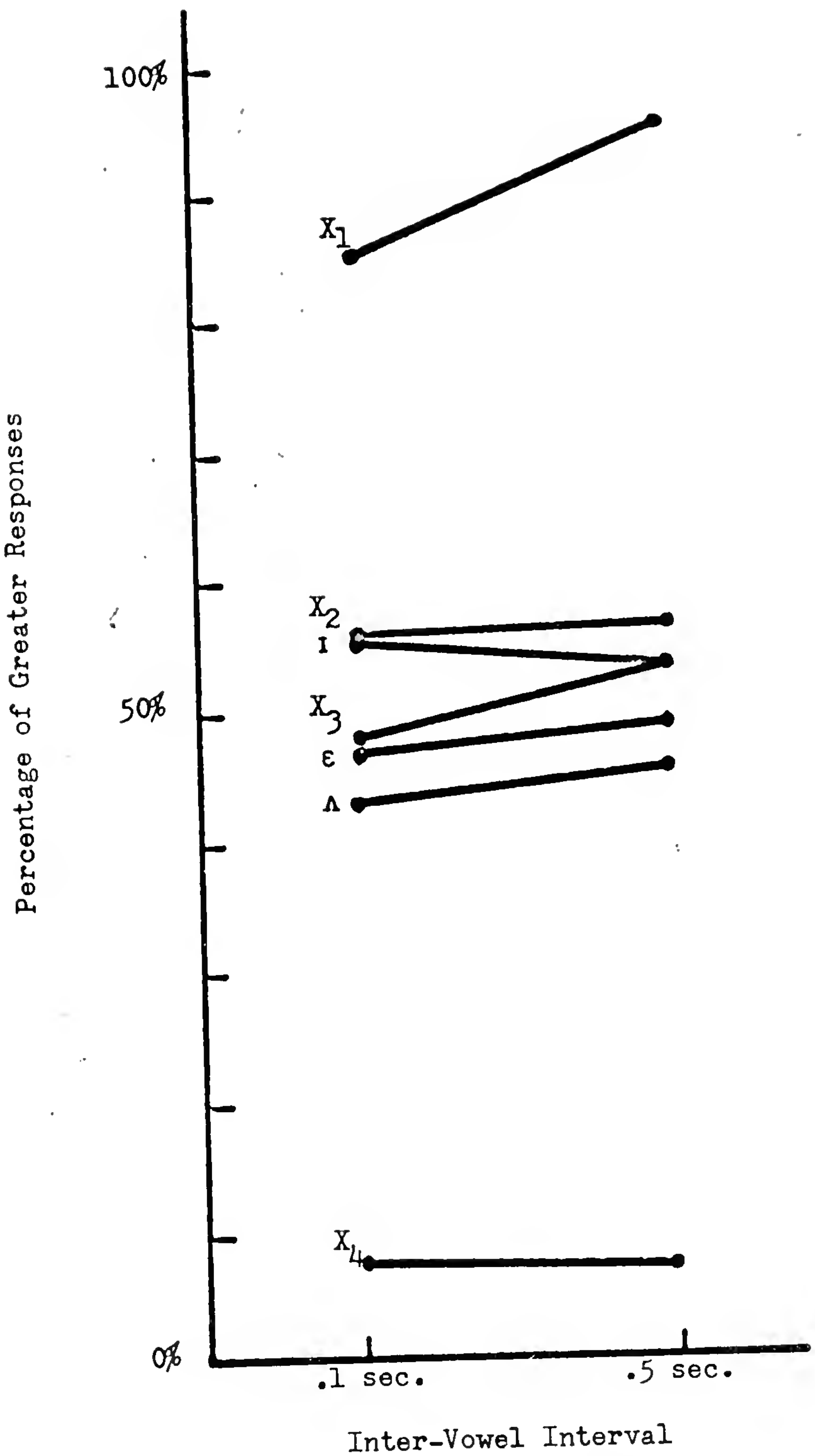

Figure 7. Percentages of greater responses for each following vowel at two inter-vowel intervals.

Table 7. Chi square values for Lv > Fv vs. Lv < Fv for each Fv.

Final Vowel	df	S of S	Chi Square
/ɪ/	1	3.760	15.05*
X_1	1	27.774	111.16*
X_2	1	22.940	91.81*
/ɛ/	1	.211	.84
X_3	1	4.214	16.86*
X_4	1	.001	.00
/ʌ/	1	.602	2.41

* = $P < .001$

In general it can be concluded that some shift in vowel identification does occur as a function of an immediately preceding vowel. The effect is of moderate strength when the two vowels are very similar and decreases as a function of the difference between the two vowels.

DISCUSSION

The results of this investigation suggest that, when two vowel samples are presented in close temporal proximity, the second will be identified as being less like the first on an F_1/F_2 continuum than would be justified by the formant values alone.

Magnitude.--The effect described above is of appreciable magnitude only when the vowel being identified (the following vowel) is relatively ambiguous. For example, the two most ambiguous vowel samples--as indicated by the proportion of responses in the modal category--were X_2 and X_3 which exhibited modal values of 56.2 and 49.1 per cent, respectively. In sharp contrast, the next highest obtained modal value was 83.4 per cent for the / ε /. When the data for these three vowels are compared with respect to the differences of greater responses, it is seen that the range for X_2 is 51.5, that for X_3 is 23.0, but that for / ε / is only 3.5. Thus, there seems to be a tendency for adjacent vowels to have their greatest effect on ambiguous vowels.

The demonstrated effect is seen also to vary as a function of the relative difference--on the F_1/F_2 dimension--between the initial and following vowels. Specifically the effect seems to be strongest when the two vowels are most similar, and to decrease in magnitude for pairs which are least similar. However, it should be noted that there were exceptions to this general relationship.

Direction.--The direction of the effect was also considered; it was found to vary as a function of the F_1/F_2 relationship of the two vowels in a pair. In cases where the F_1/F_2 ratio of the initial vowel was higher than that of the following vowel, the effect was negative; that is, fewer responses in the greater category were observed. Conversely when the F_1/F_2 ratio of the following vowel was highest, the effect tended to increase the number of greater responses. In other words, this effect is one of contrast. This finding is in agreement with the results reported by Fry (1964) and, as an effect of vowel context, it also would tend to support Joos' (1948) vowel grid hypothesis. On the other hand, these results appear to be in marked contrast to those reported by Ladefoged and Broadbent (1957) who suggested that a vowel preceded by a carrier phrase tended to be identified as the vowel within the sentence it most closely resembled. It seems possible, however, that both the findings of this study and those described by Fry could be special cases of the Ladefoged-Broadbent effect. That is, a listener hearing this type of signal may tend to place two auditorily different vowels in different phonemic categories. Thus, if he categorizes the initial vowel first, and then (on the basis of the direction of the difference of the formant one and two values for the two vowels) identifies the following vowel as being a phoneme which is in specific relationship to the initial vowel, the effect observed in this research would be obtained.

Inter-vowel interval.--The changes in the magnitude of the effect of one vowel on the identification of another as a function of inter-vowel interval actually are not clear. No consistent relationships

could be found among the data. Nevertheless, some of the data suggest that the magnitude of the contextual contrast might decrease somewhat as a function of inter-vowel interval. The inconclusiveness of this relationship, however, represents one of the questions unanswered by this research.

Vowel ambiguity.--As was noted previously the following vowels X_1 and X_4 were in reality considerably less ambiguous than were X_2 and X_3. One factor which may have contributed to this discrepancy was the possibility of unequal F_1/F_2 deviations of the four intermediate vowels from their most similar vowel. For example, re-examination of Figure 2 will reveal that the F_1/F_2 intersect for X_1 is very close to the 100 per cent (identified) area reported by Fairbanks and Grubb (1951) for the vowel / ɪ /. In marked contrast the F_1/F_2 intersect for X_2 is distant from the comparable target area of the vowel / ɛ /. A similar but less marked discrepancy is seen for the intersects of X_3 and X_4 when compared to the target areas of / ɛ / and / ʌ /. A second factor which could tend also to shift response proportions in the observed manner, was the absence of / i / and / ɑ / samples from the vowels presented for identification. If judges tended to compensate for this lack by "spreading" their responses to fill the entire continuum, a discrepancy such as that observed could be predicted. In addition, it should be noted that the major relationship found in this study, i.e., the effect of initial vowel on following vowel identification, would also act to increase this disproportionality. For both X_1 and X_2 the contextual effect would tend to inflate the number of / ɪ / responses, as three of the initial vowels would act to increase this number and only two would

tend to decrease it. Similarly for X_3 and X_4 the contextual effect would have an overall inflating effect on the number of /ʌ/ responses since three-fifths of the initial vowels would act to inflate this response category as well.

In summary, it can be concluded that the acoustic characteristics of a vowel can affect the perceptual identification of another vowel immediately following the first. This effect is one of contrast and is greatest when the two vowels are similar with respect to formants one and two.

SUMMARY AND CONCLUSIONS

The purpose of this investigation was to determine whether the identification of selected vowels is affected by the characteristics of an immediately preceding vowel. In order to accomplish this, recordings of pairs of vowels were presented to twenty-five listeners who were instructed to identify the second of each pair of stimuli.

The judged stimuli or "effected" vowels were all generated electrically. They corresponded to /ɪ /, / ɛ/ and /ʌ / or to one of four intermediate vowels. The formant frequencies of these intermediate vowels were intended to create vowels which were ambiguous; two had formant frequencies between those of /ɪ / and / ɛ/ and the other two between those of /ɛ / and /ʌ /.

Two types of vowel pairs were used. In the first type the initial or affector vowel of each pair was a human production of one of the five vowels / i/, / ɪ/, /ɛ /, / ʌ/ and /ɑ /. In the second series of pairs, two formant synthetic productions of these same five vowels were used in place of the human productions.

In order to gain information concerning whether or not the effect varies directly with the temporal spacing between the two stimuli, all possible pairings were presented with each of the two inter-vowel periods of .1 and .5 seconds.

Data were analyzed for 1) the affects of the initial vowel of each pair upon the identification of the second member of that

pair, 2) differences due to the inter-vowel periods and, 3) differences between the affects of the synthetic and human vowels.

The major conclusions provided by this research are:

1. When two vowels are presented in close temporal proximity the identification of the second is affected by the first.
2. The effect is one of contrast, that is, the second vowel of the pair is identified as though its acoustic characteristics were less similar to those of the initial vowel than would be predicted on the basis of its formant one and formant two values.
3. Vowel context has only a modest - and variable - effect on vowel identification. This effect was closely related to vowel ambiguity with the greatest effect being found for the most ambiguous samples.
4. Human and synthetic initial vowels act in a similar fashion on the identification of an immediately following vowel.
5. No conclusions may be made concerning the temporal factor as no significant relationships were found. However, there was some suggestion that the durational effects of the inter-vowel period may have been obscured by the limited temporal scale used in this study.

BIBLIOGRAPHY

Black, J. W., The nature of the spoken vowel. Arch. of Speech, 2, 7-27 (1937).

Black, J. W., Effect of consonant on the vowel. J. acoust. Soc. Amer., 10, 203-205 (1939).

DeLattre, P., Liberman, A. M., Cooper, F. S., and Gerstman, L. J., An experimental study of the acoustic determinants of vowel color. Word, 8, 195-210 (1952).

Dunn, J. K., Methods of measuring vowel formant bandwidths. J. acoust. Soc. Amer., 33, 1737-1746 (1961).

Dunn, J. K., The calculation of vowel resonances and an electrical vocal tract. J. acoust. Soc. Amer., 22, 740-753 (1950).

Fairbanks, G., and Grubb, P., A psychophysical investigation of vowel formants. J. Speech Hearing Res., 4, 203-219 (1961).

Fant, C. G. M., Transmission properties of the vocal tract with application to the acoustic specification of phonemes. Technical Report Acoustical Laboratory Massachusetts Institute Technology, 12 (January, 1952).

Fry, D. B., Experimental evidence for the phoneme, in In Honor of Daniel Jones, (D. B. Fry and D. Abercrombie, Eds.) Longmans, London (1964).

Fry, D. B., Abramson, A. S., Eimas, P. D., and Liberman, A., The identification and discrimination of synthetic vowels. Language and Speech, V, 171-189 (1962).

House, A. S., Stevens, K. N., and Fujisaki, H., Automatic measurement of the formants of vowels in diverse consonental environments. J. acoust. Soc. Amer., 32, 1517 (1960).

Joos, M., Acoustic phonetics. Language, 24, 2 (1948).

Ladefoged, F., Spectrographic determination of vowel quality. J. acoust. Soc. Amer., 32, 918-919 (1960).

Ladefoged, F., and Broadbent, D. E., Information conveyed by vowels. J. acoust.Soc. Amer., 29, 98-104 (1957).

Lewis, D., Vocal resonance. J. acoust. Soc. Amer., 8, 91-99 (1936).

Lewis, D., and Tuthill, C. E., Resonant frequencies and damping consonants of resonators in the production of sustained vowels 'O' and 'Ah'. J. acoust. Soc. Amer., 11, 451-456 (1940).

Miller, R. L., Audiology tests with synthetic vowels. J. acoust. Soc. Amer., 25, 117-121 (1953).

Peterson, G. E., The information bearing elements of speech. J. acoust. Soc. Amer., 24, 629-637 (1952).

Peterson, G. E., and Barney, H. L., Control methods used in a study of the vowels. J. acoust. Soc. Amer., 24, 175-184 (1952).

Potter, R. K., Kopp, G. A., and Green, H. C., Visible Speech, D. van Nostrand Co., Inc., N. J. (1947).

Potter, R. K., and Steinberg, J. C., Toward the specification of speech. J. acoust. Soc. Amer., 22, 807-820 (1950).

Stevens, K. N, and House, A. S., An acoustical theory of vowel production and some of its implications. J. Speech Hearing Res., 4, 303-320 (1961).

Stevens, S. S., and Davis, H., Hearing, Its Psychology and Physiology, John Wiley and Sons, Inc., N. Y. (1938).

Tiffany, W. R., Vowel recognition as a function of duration, frequency modulation and phonetic context. J. Speech Hearing Dis., 18, 289-301 (1955).

van den Berg, J., Transmission of vocal cavities. J. acoust. Soc. Amer., 27, 161-168 (1955).

APPENDIX A

Answer Sheet

1 ____	26 ____	51 ____	76 ____	101 ____	126 ____
2 ____	27 ____	52 ____	77 ____	102 ____	127 ____
3 ____	28 ____	53 ____	78 ____	103 ____	128 ____
4 ____	29 ____	54 ____	79 ____	104 ____	129 ____
5 ____	30 ____	55 ____	80 ____	105 ____	130 ____
6 ____	31 ____	56 ____	81 ____	106 ____	131 ____
7 ____	32 ____	57 ____	82 ____	107 ____	132 ____
8 ____	33 ____	58 ____	83 ____	108 ____	133 ____
9 ____	34 ____	59 ____	84 ____	109 ____	134 ____
10 ____	35 ____	60 ____	85 ____	110 ____	135 ____
11 ____	36 ____	61 ____	86 ____	111 ____	136 ____
12 ____	37 ____	62 ____	87 ____	112 ____	137 ____
13 ____	38 ____	63 ____	88 ____	113 ____	138 ____
14 ____	39 ____	64 ____	89 ____	114 ____	139 ____
15 ____	40 ____	65 ____	90 ____	115 ____	140 ____
16 ____	41 ____	66 ____	91 ____	116 ____	141 ____
17 ____	42 ____	67 ____	92 ____	117 ____	142 ____
18 ____	43 ____	68 ____	93 ____	118 ____	143 ____
19 ____	44 ____	69 ____	94 ____	119 ____	144 ____
20 ____	45 ____	70 ____	95 ____	120 ____	145 ____
21 ____	46 ____	71 ____	96 ____	121 ____	146 ____
22 ____	47 ____	72 ____	97 ____	122 ____	147 ____
23 ____	48 ____	73 ____	98 ____	123 ____	148 ____
24 ____	49 ____	74 ____	99 ____	124 ____	149 ____
25 ____	50 ____	75 ____	100 ____	125 ____	150 ____

APPENDIX B

INSTRUCTIONS TO LISTENERS

The data from these sessions are to be used to study the effect of certain variables on the perception of vowels.

Each item consists of two vowels separated by less than one second silent interval. This interval will vary.

The task is to classify the second vowel in each pair as being one of the five vowels / i /, / ɪ /, / ɛ /, / ʌ/, or / ɑ/, or if an item seems to be none of these, to decide which of them it is most like.

You will respond by filling in, in the usual electronic scoring form manner, the space corresponding to that vowel. That is, if you hear an / i / as in "peat" you would mark #1 for that item, for / ɪ / as in "pit" #2, for / ɛ / as in "pet" #3, for / ʌ/ as in "putt" #4, and for / ɑ / as in "pot" #5. Please do not leave any item blank as the scoring machine cannot score papers with blanks.

There will be 20 practice items and 280 test items. A ten second pause will follow each item. Each group of five items will be set apart by an additional five second pause. As no item numbers are given, you must be careful to follow the item numbers on the form correctly. Please note that the numbers run across the page--not down the rows. You will receive a five minute break after items 90 and 230 and a ten-fifteen minute break after item 160. There will be an opportunity to ask questions after the practice

items. If you wish to change a response, erase quickly. I will make clean erasures later.

Please do not compare responses until after the entire task is completed.

This dissertation was prepared under the direction of the chairman of the candidate's supervisory committee and has been approved by all members of that committee. It was submitted to the Dean of the College of Arts and Sciences and to the Graduate Council, and was approved as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

December 18, 1965

Ernest H. Cox

Dean, College of Arts and Sciences

Dean, Graduate School

SUPERVISORY COMMITTEE:

Chairman

Paul Moore

VITA

Carl Louis Thompson was born August 2, 1933 at Uvalde, Texas, where he received his elementary and high school education. He attended Southwest Texas Junior College during 1950-52 and received his Bachelor of Arts in 1954. He served with the United States Army from August, 1954 until August, 1956. He resumed his education at Baylor University, receiving his Master of Arts in August, 1958. He attended the University of Wichita from 1958 until 1960 and worked as an audiologist at the New Orleans Speech and Hearing Center in New Orleans, Louisiana, from July, 1960 until August, 1962. He enrolled in the Graduate School of the University of Florida in September, 1962 and worked as a graduate assistant in the Department of Speech until 1963, and remained as a predoctoral trainee until 1965. From 1962 until the present time, he has pursued his work toward the degree of Doctor of Philosophy.